VOLUME 39 MARCH 1953 NUMBER 3

PROCEEDINGS

OF THE

NATIONAL ACADEMY OF SCIENCES

OF THE

UNITED STATES OF AMERICA

EDITORIAL BOARD

LINUS PAULING, *Chairman* EDWIN B. WILSON, *Managing Editor*
ALEXANDER WETMORE, *Home Secretary* ROGER ADAMS, *Foreign Secretary*
WILLIAM W. RUBEY, *Chairman of the National Research Council*

J. W. BEAMS G. W. BEADLE GREGORY BREIT
R. E. CLELAND E. A. DOISY L. C. DUNN
W. W. RUBEY G. C. EVANS W. F. LIBBY
HARLOW SHAPLEY C. H. GRAHAM J. VON NEUMANN
F. E. TERMAN J. T. PATTERSON PAUL WEISS

Publication Office: Mack Printing Company, Easton, Pa.
Editorial Office and Home Office of the Academy
2101 Constitution Avenue, Washington 25, D. C.

Entered as Second-Class matter, June 21, 1920, at the Post-office at Easton, Pennsylvania, under the Act of August 24, 1912. Accepted for mailing at a special rate of postage provided for in section 1103, act of October 3, 1917. Authorized June 16, 1920.

Published Monthly Price per Annum, $7.50

OFFICERS OF THE ACADEMY

PRESIDENT ..... DETLEV W. BRONK
VICE PRESIDENT ..... EDWIN B. WILSON
FOREIGN SECRETARY ..... ROGER ADAMS
HOME SECRETARY ..... ALEXANDER WETMORE
TREASURER ..... WILLIAM J. ROBBINS

INFORMATION TO SUBSCRIBERS

Checks covering subscriptions to the PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES at the rate of $7.50 per annum should be made payable to the National Academy of Sciences, and sent to the Treasurer, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C. Single numbers $0.75. These prices effective January, 1951.

Microfilm copies of the PROCEEDINGS are available to subscribers to the paper edition. The microfilm is a positive, furnished on metal reels, suitably labeled, and is distributed at the end of the volume year. Inquiries concerning purchase should be directed to University Microfilms, 313 N. First St., Ann Arbor, Michigan.

Past publications of the NATIONAL ACADEMY OF SCIENCES from 1863 to 1926 and of the NATIONAL RESEARCH COUNCIL from 1916 to 1925 are listed in the Index to the First Ten Volumes of these PROCEEDINGS. This Index appeared as Part 2 of Volume 13, No. 1, January, 1927, and is for sale separately at the price of $1.50. Orders should be sent to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

Past volumes of the PROCEEDINGS may be obtained at five dollars per volume unless the sale of the volume would break a complete set of Volumes I to XXX. Single numbers may be obtained for fifty cents except where the sale of such numbers would break up a complete volume. Beginning with Volume 37 (1951), complete volumes are $7.50, and single numbers $0.75. Only forty complete sets of the PROCEEDINGS are available for sale —Volumes I to XXX price $200.00. Orders should be sent to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

THE NATIONAL ACADEMY OF SCIENCES-NATIONAL RESEARCH COUNCIL also issues technical publications reviewing and correlating recent research. The *Bulletin* and *Reprint and Circular Series* have been replaced by a single publication series, available individually, or by subscription at 25% discount. For a list of available publications, with prices and ordering procedures, address the Publications Office, National Academy of Sciences-National Research Council, 2101 Constitution Avenue, Washington 25, D. C.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 39 March 15, 1953 Number 3

NEW WHITE DWARFS IN THE SOUTHERN HEMISPHERE: THIRD LIST

By Willem J. Luyten

University of Minnesota

Communicated by Harlow Shapley, December 17, 1952

Following the color survey of proper motion stars in the Southern Hemisphere reported on previously,[1] results have now become available for the zone with declination limits $-50°$ and $-55°$. In the original survey for the detection of proper motions, 72 pairs of plates taken with the Harvard 24-inch Bruce telescope were used. These plates were blinked by the writer and some 8000 stars were found to possess appreciable motion. The motions have all been published in part C of the *General Catalogue of Proper Motions*. Additional plates with a yellow-sensitive emulsion were taken for all 72 regions with the same telescope, and these have now been compared with the original blue plates. Approximate colors for nearly all of the 8000 proper motion stars have been determined.

The vast majority of these stars are ordinary red dwarfs, of course, but some 400 were singled out as definitely bluer than expected. Narrowing down the limits still further, 63 stars have been selected from among these 400 as holding out the greatest probability of being genuine white dwarfs. Data for these stars are given in table 1, the several columns of which require no further explanation.

As before, the 63 probable white dwarfs have been arranged in four groups, as follows:

(*a*) The first group contains the 9 stars—some of which had not been published before—which had previously been identified as white dwarfs from color plates taken at the Cordoba Observatory. All of these were recoginzed again in the present survey. The last star in this group is the faint, preceding companion to Co.D. $-51°13128$, mag. 12.1 pg., spectrum MO. If this latter star is a normal main-sequence MO dwarf, the absolute magnitude of the companion would be around $+12.5$ and the star may be either an intermediate or a white dwarf.

TABLE 1

BFM	l	R.A. 1950	DEC.	m	μ	p	l.c.
		m		m	"		m
17403	228-55	3h 42.3	−51°33′	14.8	0.474	183°	+0.1
20383	250-52	10 53.3	−55 04	13.3	0.401	305	−0.3
21432	255-27	12 23.8	−53 16	13.8	0.296	298	+0.3
21970	258-46	13 23.0	−51 26	15.0	0.499	268	+0.0
23251	263-272	15 11.6	−53 47	14.7	0.319	217	+0.3
24588	269-41	16 57.8	−52 43	15.5	0.342	188	−0.2
24960	270-137	17 33.2	−54 24	16.3	0.446	193	+0.4
26377	277-60	19 48.6	−50 55	17.5	0.102	133	+0.1
27606	283-7	21 54.4	−51 14	14.7	0.410	192	+0.6
16274	219-48	0 47.8	−52 25	13.9	0.108	84	0.0
16285	220-145	48.9	−54 29	15.2	0.186	66	+0.2
16501	222-53	1 26.1	−53 17	13.8	0.204	90	0.0
17232	228-90	3 23.8	−52 34	14.9	0.042	121	+0.1
17652	230-35	4 12.1	−50 35	15.4	0.080	173	0.0
17731	230-204	4 18.1	−53 58	14.1	0.110	200	−0.1
18764	242-83	8 00.7	−53 19	15.2	0.120	76	+0.1
19001	242-13	23.8	−50 48	12.9	0.210	332	+0.1
19061	243-50	30.4	−53 30	13.8	0.167	166	0.0
19090	243-28	33.4	−52 01	13.0	0.145	106	+0.1
20912	251-24	11 21.5	−50 44	14.3	0.118	109	−0.3
21021	252-112	29.3	−53 46	14.4	0.163	275	−0.1
21065	252-73	33.5	−52 48	14.6	0.231	303	+0.3
22844	261-42	14 39.9	−52 08	15.4	0.066	218	+0.1
23823	265-18	16 00.1	−50 17	15.2	0.040	190	+0.1
24074	266-18	16 18.3	−50 45	16.1	0.168	196	+0.1
24107	266-180	20.6	−53 37	15.5	0.098	204	+0.1
24150	266-195	23.0	−54 05	15.5	0.088	16	+0.1
24378	267-102	43.1	−54 39	14.5	0.238	208	+0.1
25114	270-31	17 43.8	−52 06	15.8	0.112	148	0.0
26944	279-25	20 34.5	−53 16	14.4	0.189	162	+0.1
28016	285-14	22 48.0	−50 26	13.8	0.206	125	+0.1
16078	218-28	0 18.8	−51 57	13.9	0.062	137	+0.2
16413	221-49	1 07.2	−53 14	15.6	0.229	72	+0.4
16781	224-13	2 15.6	−50 54	15.1	0.122	170	+0.4
16907	225-64	41.8	−53 27	13.8	0.085	54	+0.2
17113	227-140	3 11.3	−54 18	14.1	0.081	163	+0.3
17132	227-121	3 13.9	−53 41	13.0	0.134	194	+0.3
17257	228-62	26.6	−51 44	16.8	0.238	78	+0.5
17557	230-63	4 04.1	−51 04	14.9	0.106	86	+0.2
18208	237-71	6 23.7	−52 59	15.2	0.087	95	+0.2
18272	237-54	32.7	−52 18	14.2	0.050	165	+0.2
19844	247-17	10 04.4	−50 55	12.0	0.138	201	+0.2
20110	248-43	28.0	−53 17	15.1	0.057	250	+0.1
20160	249-28	32.9	−51 33	14.2	0.066	331	+0.2
20364	249-89	52.0	−53 48	12.9	0.127	143	+0.2
20933	251-194	11 22.7	−53 06	13.4	0.063	270	+0.1

TABLE 1 (Continued)

BPM		R.A. 1950	DEC.	m	μ	p	I.C.
		m		m	"		m
21487	255-18	12 32.3	−51 27	15.1	0.165	261	+0.3
22719	260-46	14 27.1	−51 47	14.9	0.056	293	+0.2
22878	262-9	43.6	−50 47	14.0	0.063	247	+0.2
23075	262-18	15 00.4	−51 14	13.4	0.063	241	+0.2
23380	263-9	19.6	−50 20	14.0	0.079	232	+0.2
23943	265-55	16 08.3	−50 58	15.6	0.064	236	+0.3
24123	266-22	21.2	−51 09	15.6	0.093	223	+0.2
24163	266-64	23.6	−51 54	15.2	0.075	263	+0.2
24383	268-31	43.7	−51 20	14.3	0.053	297	+0.2
24601	268-92	59.0	−53 06	13.6	0.134	175	+0.1
26631	277-125	20 08.0	−52 21	11.8	0.212	190	+0.3
28452	289-12	23 46.3	−50 43	14.5	0.061	169	+0.3
24203	266-168	16 27.1	−53 31	16.0	0.067	244	<+0.2
24754	269-72	17 14.8	−54 44	16.2	0.164	151	<+0.2
25260	271-168	54.5	−55 01	15.4	0.130	219	<+0.2
26456	277-155	19 54.9	−53 02	17.8	0.091	170	<+0.4
27369	282-25	21 27.3	−51 00	15.9	0.085	122	<+0.4

[a] In *Ap. J.*, **96,** 58 (1942), the color of this star was given as +1.0, determined from plates taken at Cordoba. This, however, is an error; the reddish color belongs to BPM 16273.

(*b*) The second group comprises the 22 stars for which the probability that they are white dwarfs is very high.

(*c*) The third group contains 27 stars, some of which will almost certainly prove not to be white dwarfs. But while this will remove several stars from the list, others from among the 400 are expected to take their place, and, statistically speaking, the stars in this group should be added to the number in group (*b*).

(*d*) The fourth group contains those 5 stars which were invisible on the yellow plates. Past experience has shown that the chance of their being really white is very high but not much more can be said at the present time.

If all the stars announced in the present list and in the two lists previously published are counted as probable white dwarfs, then the total number of such stars now known has reached 267.

Acknowledgment.—It is a pleasure to acknowledge the financial support received from the Rumford Fund of the American Academy of Arts and Sciences which made possible the taking of the yellow plates. The author is further indebted to the Harvard College Observatory, and especially to Dr. Harlow Shapley for his willingness to have these plates taken and for his continued interest in and support of this program of research.

[1] These PROCEEDINGS, **37,** 637–640 (1951) and **38,** 494–496 (1952).

A KINETIC TREATMENT OF SATURATION AND CONDENSATION STATISTICS

BY HENRY EYRING AND MERRILL WALLENSTEIN

UNIVERSITY OF UTAH, SALT LAKE CITY

Communicated December 26, 1952

Many of the processes in chemistry and physics either involve or can be considered to involve the escape of an atom, molecule, or elementary particle from one set of states to another. As real examples, we may mention the escape of a molecule or atom from the gas phase to a liquid phase, or from a gas or liquid phase to the surface of an adsorbent. On the other hand, one can view the problem of distribution of electrons among a set of degenerate quantum states as essentially a rate process in which electrons escape from some volume of otherwise free space into these degenerate states. Among such processes it is possible to identify certain classes which have properties in common. In the case of gases adsorbing on a solid, two distinct classes are defined by the chemical adsorption of a monolayer and the physical adsorption of a multiplicity of layers. In the first case, a molecule attaches itself to a surface site and "saturates" it in the sense that no other molecule may simultaneously attach itself to the same site; while in the second case, the individual adsorption sites can accommodate an unlimited number of molecules which are, in effect, piled one atop the other on the same surface site. This situation is paralleled in statistical mechanics by systems obeying Fermi-Dirac statistics on the one hand and by systems conforming to the requirements of Bose-Einstein statistics on the other. In the case of Fermi-Dirac statistics, the adsorption sites become the members of a set of degenerate states and the adsorbing units may now be any elementary particle of odd parity. Here, considering electrons as an example, the act of putting an electron in one of the set of states cancels any further interaction of that state with other electrons while the first electron is still in occupation (Pauli principle). We may say that this "site" is saturated in precisely the same sense in which the molecule saturated a surface site in chemical adsorption. Bose-Einstein statistics, which applies to particles of even parity, makes no restriction on the number of particles which can be placed in a given quantum state and thus corresponds closely to the phenomenon of multilayer adsorption. Just how closely these apparently unrelated phenomena are connected can be brought out quite clearly by a rather simple mathematical formulation, the form of which is due originally to van't Hoff and which was used by Langmuir to derive his well-known adsorption isotherm.

GENERAL THEORY

For generality we shall refer to a number, n, of units which may, in various particular cases, be atoms, molecules or elementary particles and a number of sites which may be adsorption sites on a surface, or a set of quantum states, also depending on the phenomenon under consideration. Let the number of these sites be ω_i and the energy of interaction between any one of these sites and a unit be ϵ_i. This quantity in the ideal case will be constant for all sites being considered, but in other cases, it may vary from one site to another. This will not affect the general theory but, in application, will necessitate a knowledge of ϵ_i as a function of coverage. We shall consider the instant when n_i units are adsorbed on the ω_i sites leaving n_g free units in some volume V. This volume will, for the present, be considered large enough so that the free units possess the properties of an ideal gas. At any instant the n_i adsorbed units are leaving the sites at a velocity $n_i k_i'$ where k_i' is the specific rate constant for the dissociation of a site and an adsorbed unit. The reverse process in which free units go over to the bound state requires some consideration. If the process is one of saturation, then clearly, the instantaneous number of unsaturated sites is $(\omega_i - n_i)$ and the rate at which units disappear from the free state is $k_g' n_g(\omega_i - n_i)$, where again k_g' is the specific rate constant. At equilibrium the two rates are equal; hence

$$k_i' n_i = (\omega_i - n_i) n_g k_g'. \tag{1}$$

This is the general equation for any phenomenon involving saturation.

For the case where a site can accommodate an unlimited number of units, the expression corresponding to (1) may be derived by appealing to the principle of detailed balance.[1] Whereas, in deriving equation (1), the situation could be pictured as in figure 1 in which a unit, O, completely saturates a site, ×, the case here is as indicated in figure 2. As before, any one of the n_i units is a candidate for removal to the free state. If at any time a unit such as ● in figure 2, leaves its bound position, then

FIGURE 1 FIGURE 2

according to the principle of detailed balance there must be a reverse process which replaces this unit. Thus, *in effect*, the crosses in figure 2 indicate possible sites for accommodation of a gas molecule so that, at

any instant, the number of effective sites for the accommodation of a free unit is $(\omega_i + n_i)$ and at equilibrium

$$k_i' n_i = (\omega_i + n_i) n_g k_g'. \tag{2}$$

Equation (2) can be considered as the general equation for any process in which a given site can accommodate an unlimited number of units. It will be referred to as the general equation of condensation statistics.

In either the case of saturation or condensation statistics it may be that for a given situation $n_i \ll \omega_i$ at equilibrium, so that both (1) and (2) reduce to

$$k_i' n_i = \omega_i n_g k_g'. \tag{3}$$

Such a relation obviously corresponds to the reduction of quantum statistics to classical statistics. This form is also obtained if we consider the

FIGURE 3

case illustrated by figure 3 in which a unit interacting with a site neither uses up the site nor creates a new one as in the previous cases.

Let us next consider the case where a site is saturated by d units. Then in place of equation (2), we have

$$k_i' n_i = (\omega_i + n_i) n_g k_g' - p(d + 1)\omega_i n_g k_g', \tag{4}$$

where p is the probability that any one of the ω_i sites has become saturated and thus deletes $(d + 1)$ condensation sites from the set $(\omega_i + n_i)$. In the special case in which d is infinite, p reduces to zero whence equation (4) reduces to (2). When d is unity then $p = n_i/\omega_i$ and equation (4) reduces to equation (1), as it must. Thus, equation (4) is the general result for all types of condensation but the values of p are not of such simple form for other values of d. Equation (4), with a suitable expression for p, is appropriate, for example, for treating condensation in a capillary d units in width. Gentile[2] has treated certain aspects of this problem.

Applications

Let us now consider equations (1) and (2) together in the form

$$k_i' n_i = (\omega_i \pm n_i) n_g k_g'. \tag{5}$$

According to absolute reaction rate theory, the specific rate of any unit process can be written

$$k' = \frac{\sum_i e^{-\epsilon_i/kT} \nu_i \gamma_i + \kappa' \frac{kT}{h} F^{\ddagger} e^{-\epsilon^{\ddagger}/kT}}{F} . \tag{6a}$$

Here F is the partition function for molecules of the reacting species entering into the activated complex and $F^{\ddagger}$ that for the activated complex. $\epsilon^{\ddagger}$ is the energy of activation at absolute zero. The first term in the numerator of (6) takes care of the possibility of leakage through the potential barrier between reactants and products. In this term, ϵ_i is the energy of the system in the ith state ($\epsilon_i < \epsilon^{\ddagger}$), ν_i is the frequency of vibration in the ith state and γ_i is the chance per vibration of passing from the initial to the activated state. κ' is the transmission coefficient for complexes with energies as large or larger than $\epsilon^{\ddagger}$. Under most circumstances, we simply write

$$k' = \kappa \frac{kT}{h} \frac{F^{\ddagger}}{F} e^{-\epsilon^{\ddagger}/kT} \tag{6b}$$

which defines κ, the complete transmission coefficient. For many systems of interest κ can be taken as unity. Substituting (6) in (5) and cancelling similar terms on either side yields

$$\frac{n_i}{F_{gs}} e^{-\epsilon^{\ddagger}_{gs(i)}/kT} = \frac{(\omega_i \pm n_i) n_g}{F_g F_s} e^{-\epsilon^{\ddagger}_{g+s(i)}/kT} . \tag{7}$$

The subscripts g and s refer to the free units and sites, respectively, and gs refers to the combined unit and site in the complex or the products. Writing $\epsilon_i = \epsilon^{\ddagger}_{g+s(i)} - \epsilon^{\ddagger}_{gs(i)}$, which is just the energy of the process represented by (5) at absolute zero, and solving for n_i yields

$$n_i = \frac{\omega_i}{\frac{F_g F_s}{F_{gs} n_g} e^{\epsilon_i/kT} \pm 1} . \tag{8}$$

Under the assumption that the free units have the properties of an ideal gas, the partition function per molecule is

$$F_g = \frac{(2\pi mkT)^{3/2}}{h^3} V j^{(T)} \tag{9}$$

where $j^{(T)}$ represents the internal degree of freedom of the free units. The free energy of the n_g free units in the volume V is then

$$F = n_g kT \ln \frac{(2\pi mkT)^{3/2}}{h^3} \frac{V}{n_g} j^{(T)} \tag{10}$$

$$= -n_g kT \ln \frac{F_g}{n_g} . \tag{11}$$

The chemical potential, μ, per molecule then satisfies the equation

$$\frac{\mu}{kT} = -\ln\frac{F_g}{n_g} \text{ and } \frac{F_g}{n_g} = e^{-\mu/kT}. \tag{12}$$

This result in (8) gives

$$n_i = \frac{\omega_i}{\frac{F_s}{F_{gs}}e^{(\epsilon_i-\mu)/kT} \pm 1}. \tag{13}$$

From this equation, we may begin to consider special cases. For electrons, it is clear that F_s and F_{gs} are both unity so that

$$n_i = \frac{\omega_i}{e^{(\epsilon_i-\mu)/kT}+1}. \tag{14}$$

This is just the equation of Fermi-Dirac statistics. Similarly, for any particles of even parity, we have, for the same reasons, the equation for Bose-Einstein statistics:

$$n_i = \frac{\omega_i}{e^{(\epsilon_i-\mu)/kT}-1}. \tag{15}$$

Frenkel[3] has used a kinetic approach to derive Bose-Einstein and Fermi-Dirac statistics.

Statistical Thermodynamics.—From thermodynamics, we have for the average energy, E, of a system, the expression

$$E = -T^2\left[\frac{\partial}{\partial T}\left(\frac{A}{T}\right)\right]_V. \tag{16}$$

Here, A is the maximum work function or Helmholtz free energy. We can also write

$$E = \frac{\sum_i \epsilon_i n_i}{\sum_i n_i}. \tag{17}$$

Thus,

$$-T^2\left[\frac{\partial}{\partial T}\left(\frac{A}{T}\right)\right]_V = \frac{\sum_i \epsilon_i n_i}{\sum_i n_i}. \tag{18}$$

Now it is well known that if A is known as a function of T and V, all thermodynamic properties can be calculated. This general result takes a simple form for systems obeying classical statistics. Substituting equation (13) in (18) and remembering that in the classical case unity may be neglected in comparison with $\frac{F_s}{F_{gs}}e^{(\epsilon_i-\mu)/kT}$ we obtain

$$-T^2\left[\frac{\partial}{\partial T}\left(\frac{A}{T}\right)\right]_V = \frac{\sum_i \epsilon_i \omega_i e^{-\epsilon_i/kT}}{\sum_i \omega_i e^{-\epsilon_i/kT}}$$

$$\equiv kT^2 \frac{\partial}{\partial T} \ln \sum_i \omega_i e^{-\epsilon_i/kT}. \tag{19}$$

Whence

$$A = -kT \ln \sum_i \omega_i e^{-\epsilon_i/kT} + cT. \tag{20}$$

Now if the entropy is to be zero at the absolute zero in accordance with the third law, we must have $\omega_i = 1$ and $c = 0$. Thus, for classical statistics, we obtain the familiar result

$$A = kT \ln \sum_i' \omega_i e^{-\epsilon_i/kT}. \tag{21}$$

The Langmuir Adsorption Isotherm and Multilayer Condensation.—In equation (8) we write $n_i/\omega_i = \sigma$ as the fraction of the surface covered and also write

$$\frac{F_g F_s}{F_{gs} n_g} e^{\epsilon_i/kT} = \frac{F_g F_s}{F_{gs}\left(\frac{pV}{kT}\right)} e^{\epsilon_i/kT} = \frac{k}{p}, \tag{22}$$

so that equation (8) becomes

$$\sigma = \frac{1}{\frac{k}{p} \pm 1} = \frac{p}{k \pm p} \tag{23}$$

whence

$$p = k \frac{\sigma}{1 \pm \sigma}. \tag{24}$$

the plus sign in (23) and the minus sign in (24) correspond to Langmuir's monolayer adsorption isotherm for a uniform surface.

The choice of the $-$ sign in (23) gives

$$\sigma = \frac{p}{k - p} \tag{25}$$

and corresponds to multilayer adsorption with equal binding energy between layers. This is approximately fulfilled for layers condensing on a liquid surface. It is interesting that (25) predicts condensation in a closed system under certain circumstances. For example, if p starts out larger than K, the system is supersaturated. Equilibrium can be established only after a sufficient number of layers, as measured by σ in suitable units, have condensed so that the pressure p is reduced to the point where equation (25) is satisfied.

That one should not be surprised at this result can be seen from the following discussion. Consider equation (13) with the negative sign, viz.:

$$n_i = \frac{\omega_i}{\frac{F_{gs}}{F_s} e^{(\epsilon_i - \mu)/kT} - 1} \tag{13a}$$

When $F_{gs}/F_s = 1$, this equation yields the Bose-Einstein result. It is well known that a perfect Bose-Einstein gas displays a phenomenon closely analogous to classical condensation.[7] The discussion of condensation in a Bose-Einstein gas leads to an equation of just the same form as that from which the conditions for condensation (derived from the cluster theory of Mayer[9]) in a classical gas are deduced.[8, 9] Thus, one anticipates the relation (13a) which bears essentially the same relations to classical condensation as the Bose-Einstein equation bears to statistical condensation.

The Richardson Equation.—The number of electrons striking a square centimeter of surface per second, according to the rate equation, is

$$\begin{aligned} n_g k' &= \frac{n_g \kappa \frac{kT}{h} \frac{2\pi mkT}{h^2} \cdot 2 \cdot e^{-\epsilon^{\pm}/kT}}{\frac{(2\pi mkT)^{3/2}}{h^3} \cdot 2V} \\ &= \kappa \frac{kT}{h} \frac{2\pi mkT}{h^2} \cdot 2 \cdot e^{-(\epsilon^{\pm} + W)/kT} \end{aligned} \tag{26}$$

where the factor 2 takes account of the spin degeneracy of the electron and where

$$W = kT \ln \frac{(2\pi mkT)^{3/2}}{h^3} \cdot 2 \cdot \frac{V}{n_g} \tag{27}$$

Here, W is called the work function and is the negative of the chemical potential of the electrons in the gas phase and, therefore, also of the electrons in the metal which is in equilibrium with the gas. The Richardson's equation[4, 5]

$$I = en_g k' = \kappa \frac{4\pi em(kT)^2}{h^3} e^{-(\epsilon^{\pm} + W)/kT} \tag{28}$$

for the current leaving a metal surface agrees with experiment if $\epsilon^{\pm}$ be taken equal to zero and the transmission coefficient, κ, which in this case, is the accommodation coefficient, is taken equal to one-half. Equation (18) permits one to calculate the chemical potential and, therefore, n_g directly. The energy levels in a metal are ordinarily not well enough known, however, to justify this direct approach.

Saturation Phenomena in Homogeneous Systems.—Saturation phenomena are likewise familiar in homogeneous catalysis. The intensity, I, of luminescent bacteria may be expressed by the relation

$$I = nk'(LH_2)(A_n) \tag{29}$$

Here A_n is the concentration of the active enzyme luciferase; (LH_2) is the concentration of the luciferin which is being oxidized; k' is the specific rate constant and b is a proportionality factor. As the temperature is raised, active luciferin passes reversibly into an inactive form A_d. Thus, we have the equilibrium relationship

$$\frac{(A_d)}{(A_n)} = K \tag{30}$$

which, together with the conservation equation for total luciferase,

$$(A_0) = (A_n) + (A_d) = (A_n) + K(A_n),$$

gives

$$(A_n) = \frac{(A_0)}{1 + K} \tag{31}$$

and finally the result[6]

$$I = \frac{bk'}{1 + K}(LH_2)(A_0). \tag{32}$$

If one considers further the effect of enzyme inhibitors, one, of course, is led to equations equivalent to those describing the effect of poisons on surfaces in the Langmuir Theory.

Acknowledgment.—The authors wish to express their thanks to the National Science Foundation for their part in supporting this work.

[1] Tolman, R. C., *The Principles of Statistical Mechanics*, Oxford University Press, 1938.

[2] Gentile, G., *Nuov. Cim.*, **17**, 493 (1940). For further references and discussions see Partington, J. R., *Advanced Treatise of Physical Chemistry*, Vol. I, Longmans, 1949.

[3] Frenkel, J., *Wave Mechanics*, 2nd ed., Oxford University Press, 1936.

[4] von Laue, M., *Jahrb. Radioact. Electronik*, **15**, 205 (1918).

[5] Dushman, S., *Phys. Rev.*, **21**, 63 (1923). See, for example, Mayer and Mayer, *Statistical Mechanics*, John Wiley & Sons, 1940, for a more detailed discussion.

[6] For an elaboration of a wide variety of types of saturation of enzymes in homogeneous systems, reference may be made to Johnson, F. H., Eyring, H. and Polissar, M. J., *The Kinetic Basis of Molecular Biology*, John Wiley & Sons, Inc.; (in press).

[7] Einstein, A., *Ber. Berl. Akad.*, 261 (1924) and 3 (1925).

[8] Kahn, B., and Uhlenbeck, G. E., *Physica*, I, 399 (1938).

[9] See also, Mayer and Mayer, *Statistical Mechanics*, John Wiley & Sons, Inc., 1940, for detailed discussion.

THE HIGH FIELD CONDUCTANCE OF SOME SYMMETRICAL VALENCE-TYPE STRONG ELECTROLYTES*

By Andrew Patterson, Jr.

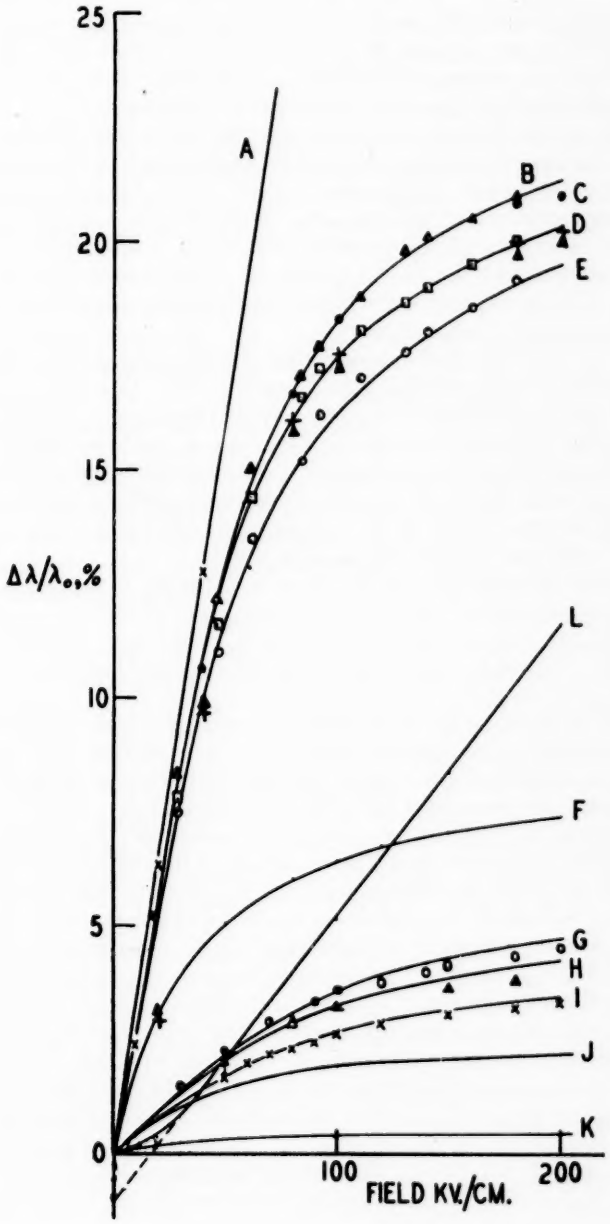
Sterling Chemistry Laboratory, Yale University, New Haven, Connecticut†

Communicated by H. S. Harned, January 21, 1953; read before the Academy November 7, 1951

In 1927 Wien[1] announced the significant discovery that the conductances of electrolytic solutions do not conform to Ohm's law under the influence of high potential gradients. In the years which followed, up to about 1939, Wien and his coworkers established the nature of the deviation from Ohm's law as an increase of conductance with increasing field strength, with quite different effects for weak and for strong electrolytes. This work has been summarized by Eckstrom and Schmelzer.[2] Theoretical explanations for the so-called Wien effect were devised by Onsager for weak electrolytes[3] and by Onsager and Wilson for symmetrical valence-type strong electrolytes.[4] The nature of the conductance increase is shown in figure 1, where the fractional conductance increase (in per cent) is plotted as a function of the field strength; curve *K* is a theoretical computation[4] for potassium chloride; curves *G*, *H*, and *I* are experimental data for copper, zinc, and magnesium sulfates, respectively; curves *B*, *D*, and *E* are for lanthanum ferricyanide at three slightly different concentrations. All these compounds are strong electrolytes. Curve *L* is for a typical weak electrolyte, acetic acid. The data for all these electrolytes pertain to solutions close to 10^{-4} molar at 25°. As is clearly shown in figure 1, all the strong elec-

FIGURE 1

The high field conductance of some symmetrical valence-type strong electrolytes: *A*, theoretical corrected Onsager computation for lanthanum ferricyanide using $c = 1.025 \times 10^{-4}$ molar and $K(0) = 1.82 \times 10^{-4}$; *B*, *D*, and *E*, experimental results for lanthanum ferricyanide, 1.032×10^{-4} molar (△), 1.025×10^{-4} molar (□), and 1.023×10^{-4} molar (○); *C*, ●, corrected theoretical Onsager computation using $c = 1.025 \times 10^{-4}$ molar and $K(0) = 3.50 \times 10^{-4}$; *D*, +, corrected theoretical Onsager computation using $c = 1.032 \times 10^{-4}$ molar and $K(0) = 3.702 \times 10^{-4}$; *D*, ▲, corrected theoretical Onsager computation using $c = 1.025 \times 10^{-4}$ molar and $K(0) = 3.702 \times 10^{-4}$; *F*, theoretical uncorrected Onsager-Wilson computation for lanthanum ferricyanide, 1.025×10^{-4} molar; *G*, corrected theoretical Onsager computation (solid curve) and experiment (○) for copper sulfate, 1.77×10^{-4} molar; *H*, corrected theoretical Onsager computation (solid curve) and experiment (△) for zinc sulfate, 1.64×10^{-4} molar; *I*, corrected theoretical Onsager computation (solid curve) and experiment (×) for magnesium sulfate, 1.39×10^{-4} molar; *J*, uncorrected Onsager-Wilson theoretical computation for magnesium sulfate, 1.39×10^{-4} molar; *K*, theoretical Onsager-Wilson computation for potassium chloride, 2.335×10^{-4} molar; *L*, acetic acid, 7.405×10^{-4} molar; all results and computations are for a temperature of 25.00°.


25
20
15
10
5
0
Δλ/λ₀,%
100
200
FIELD KV./CM.
A
B
C
D
E
F
G
H
I
J
K
L

trolytes exhibit a rapid increase of conductance at comparatively low fields and a much slower increase at higher fields; weak electrolytes show an apparently linear increase of conductance with field. With the 3–3 electrolyte the fractional increase of conductance is quite large.

Recently, Gledhill and Patterson[5] have devised a new and improved method for the measurement of high field conductance; as a result a considerable body of data has become available under circumstances and of sufficient precision such that meaningful quantitative relations may be drawn between the high field conductance behavior of different compounds, a situation which has not existed heretofore. In the present paper we shall discuss relations between the high field conductances of symmetrical 2–2 and 3–3 valence-type electrolytes.

The experimental results presented for discussion are given in figure 1. Curve G is for copper sulfate, 1.77×10^{-4} molar.[6] Curve H is for zinc sulfate, 1.64×10^{-4} molar.[7] Curve I is for magnesium sulfate, 1.39×10^{-4} molar.[7] In each case the symbols refer to the experimental data, while the solid lines refer to theoretical computations to be discussed below. Curves B, D, and E are for lanthanum ferricyanide at concentrations 1.032×10^{-4} molar, 1.025×10^{-4} molar, and 1.023×10^{-4} molar, respectively.[8] All measurements were made at 25.00°.

Through the use of the Onsager-Wilson modification of the Onsager conductance equation,[4]

$$\Lambda_x = \Lambda^0 - \left(\frac{3}{2 - \sqrt{2}} \alpha^* g(x) \Lambda^0 + \frac{\beta^*}{\sqrt{2}} f(x) \right) c^{1/2} \qquad (1)$$

with the terms as defined by Harned and Owen,[4] it is possible to compute theoretical curves for these electrolytes. A curve derived from equation (1) has been drawn for magnesium sulfate, curve J, and a similarly computed curve has been drawn for lanthanum ferricyanide, $c = 1.025 \times 10^{-4}$ molar, curve F. Both of these curves fall considerably below the experimental results to which they should correspond, the difference being particularly bad in the case of the 3–3 electrolyte. A similar disparity is noted between theoretically computed and experimental values for the low field conductance of 2–2 and 3–3 electrolytes; on the other hand, 1–1 strong electrolytes conform nicely to theory in conductance, transference numbers, and other properties. A satisfactory solution to the problem is found in the assumption of incomplete dissociation into ions and the assignment of an equilibrium constant to express the degree of "weakness" of the higher valence-type strong electrolytes; one of the most probable causes for this weakness is the formation of ion-pairs by the higher valence-type ions under the influence of coulombic forces. Bailey and Patterson[7] found it possible greatly to minimize the difference between experiment and theory for the high field conductance of 2–2 electrolytes, as will be seen from an

inspection of the solid lines (theory, corrected) and the experimental values for copper, zinc, and magnesium sulfates (curves *G*, *H*, and *I*), as compared with the uncorrected Onsager theory for magnesium sulfate, curve *J*. This improvement was accomplished by combining the Onsager theories for weak and strong electrolytes, on the assumption that the ion-pairs present obeyed the Onsager theory for weak electrolytes, while the ions obeyed the Onsager-Wilson theory for strong electrolytes. The net result is a modification of equation (1) to yield the similar equation

$$\Lambda_x = \alpha \left\{ \Lambda^0 - \left(\frac{3}{2 - \sqrt{2}} \alpha^* \Lambda^0 g(x) + \frac{\beta^*}{\sqrt{2}} f(x) \right) (\alpha c)^{1/2} \right\} \qquad (2)$$

in which the degree of dissociation, α, is a function of field and x, the argument of Onsager and Wilson, is a function of α.

It is of interest to note that the high field conductance of the three 2–2 electrolytes computed with equation (2) in each case falls slightly above the experimental results, although significantly closer than curves computed from equation (1) of which *J* is representative. The degree of dissociation is obtained from the equilibrium constant at zero field, obtained from low field conductance measurements on the three electrolytes. The position of the corrected curves is principally dependent upon the value of $K(0)$. To derive a value of $K(0)$ from a set of conductance data it is necessary to know what the conductance behavior of an ideal non-associated 2–2 electrolyte would be in order to have a basis for comparison; this is a practical requirement for the extrapolation of the data. Also, in conductance measurements corrections must be made for the conductances of hydrolysis products, for example, MOH^+ and HSO_4^- in the case of the sulfates under consideration. As a result, the precision with which an equilibrium constant value can be determined from conductance data is not high. It is thus surprising that the results obtained from equation (2) are as good as they are, but it is more surprising that each theoretical curve is just slightly higher than the experiment at the same field. This difference is in the direction of requiring a somewhat larger value of $K(0)$ for best agreement.

When equation (2) is applied to the data for lanthanum ferricyanide the theoretical values obtained fall far above experiment; at 200 kv./cm., for example, the theoretical value is 36.3, compared to curve *D*, 20.0; see curves *A* and *D*, figure 1. With the 3–3 electrolyte α_0 is naturally smaller since the electrolyte is weaker or more associated, and consequently the change of α with increasing field is greater and of more importance in determining the position of the theoretical curve. As before, $\Delta\lambda/\lambda_0$ is a sensitive function of $K(0)$. The fractional high field conductance quotient, $\Delta\lambda/\lambda_0$, is a sensitive function of concentration also, with this high valence-type electrolyte, as will be observed from curves *B*, *D*, and *E*; the variation

is consistent with the usual change of $\Delta\lambda/\lambda_0$ with concentration: the higher the concentration, the larger the fractional conductance quotient.

Because of the difficulties attendant upon a determination of $K(0)$ from conductance data for this high valence-type electrolyte we have employed our high field conductance data to obtain a value of $K(0)$ which is consistent with these data. This appears to be justified from a number of points of view, and it is possible that a more meaningful value than obtained from low field conductance work may result. As is noted in reference 8, this requires a series of successive approximations, since the values of most of the quantities relating $K(0)$ to α_0, and thus to the high field conductance quotient, are interrelated. The result obtained from this procedure is a consistent set of values of α_0 and $K(0)$. For example, for curve D, α_0 has the value 0.87900 and $K(0)$ is 3.702×10^{-4}, compared with the value 1.82×10^{-4} obtained by Davies and James.[9] These quantities differ by a factor of two.

Several comments may be made on this procedure for obtaining a value for $K(0)$. It is assumed that $\gamma_\pm$ is unity throughout the range of applied fields when equation (2) is employed; at zero field, a value is obtained from other measurements. As the field is increased $\gamma_\pm$ approaches unity; at a field of 200 kv./cm. for which the computation of $K(0)$ has been made the assumption should not have much influence, but at low fields the theoretical curve falls noticeably below the experimental values. It is to be noted that the assumption that $\gamma_\pm$ is unity would cause the theoretical curves to fall slightly higher than the experimental values, which is indeed the case with the 2–2 electrolytes examined. This assumption would play a larger part in the calculations for a 3–3 electrolyte since the degree of association is larger than with 2–2 electrolytes. On the other hand, the presence of impurities or hydrolysis products will have much less effect on the relative conductance increase than on an absolute conductance measurement, especially if the hydrolysis products or impurities be of lower valence type than the electrolyte under study. Also, except for the difficulty of the assumption that $\gamma_\pm = 1$, no assumptions are made *per se* as to the behavior of a perfect 3–3 electrolyte. It is assumed that Onsager's theory for weak electrolyte high field conductance is valid; this fortunately is a matter which is well established.

Although the employment of equation (2) involves a combination of both Onsager theories for weak and for strong electrolytes, it is important to recall that the Onsager theory for weak electrolytes makes no attempt to include the effect of the field on the free ions present or formed by the field; since the measurements which give the results in figure 1 are relative to potassium chloride as a reference electrolyte, the effect of the strong electrolyte behavior of the free ions produced by the field is compensated for very effectively if they are univalent, and somewhat less so if of higher valence type.

The consistent values of $K(0) = 3.702 \times 10^{-4}$ and $\alpha_0 = 0.87900$ when employed with the concentration $c = 1.025 \times 10^{-4}$ molar give rise to the points rendered as filled triangles near curve D. The agreement is very satisfactory since the data are derived one from the other. When a value of $K(0) = 3.50 \times 10^{-4}$ is used with the concentration 1.025×10^{-4} molar, the filled circles marked curve C result. This curve deviates from curve D by about one per cent in the term $\Delta\lambda/\lambda_0$ at 200 kv./cm. Since the precision of measurement of the fractional high field conductance quotient is in the order of 0.03, this is then a sensitive method of obtaining a value for $K(0)$. Since $K(0)$ is supposed to be a constant, the value 3.7×10^{-4} was employed with the concentration 1.032×10^{-4} molar, for which the data of curve B were obtained experimentally. The result was the symbols drawn as plus marks near curve D. Although these data lie in the correct direction from the filled triangles (that is, higher), they fall considerably below the experimental results, curve B. The best explanation which can be offered for this lack of agreement is that the expression

$$\log \frac{\alpha^2 c}{3(1-\alpha)} + \log \frac{y_\pm^2}{y_u} = \log K(0) \tag{3}$$

where

$$\log \frac{y_\pm^2}{y_u} = -9 \left\{ \frac{\Gamma^{1/2}}{1 + \Gamma^{1/2}} - 0.20\Gamma \right\} \tag{4}$$

(with all terms as defined by Harned and Owen[4]) must be employed to relate α_0 and $K(0)$. Equations (3) and (4) involve an empirical extension of the Debye-Hückel-Onsager limiting equation which has been extensively used with other electrolytes of lower valence type. Only very limited data are available to test the equation with 3–3 electrolytes. If the ratio, $y_\pm/y_u$, calculated for the larger concentration (1.032×10^{-4}) were larger than it should be, and if the same ratio calculated for the smaller concentration (1.025×10^{-4}) were more nearly correct, then the calculated points would be lower than they should be for the higher concentration. It is obviously desirable to extend the measurements to a wider range of concentrations; it would also be most helpful to have more experience with the use of equations (3) and (4) for 3–3 electrolytes.

* Presented in large part before the Autumn meeting of the National Academy of Sciences at Yale University, November 7, 1951. This paper reports work done by Dr. Frederick E. Bailey, Jr., Mr. Daniel Berg, and the author.

† Contribution No. 1139 from the Department of Chemistry, Yale University.

[1] Wien, M., and Malsch, J., *Ann. Physik*, **83**, 305 (1927).

[2] Eckstrom, H. C., and Schmelzer, C., *Chem. Revs.*, **24**, 367 (1939).

[3] Onsager, L., *J. Chem. Phys.*, **2**, 599 (1934).

[4] Harned, H. S., and Owen, B. B., *The Physical Chemistry of Electrolytic Solutions*, 2nd ed. Reinhold Publishing Corp., New York, N. Y., 1950, pp. 95–114.

[5] Gledhill, J. A., and Patterson, A., *J. Phys. Chem.*, **56**, 999 (1952).
[6] Berg, D., and Patterson, A., *J. Am. Chem. Soc.*, **74**, 4704 (1952).
[7] Bailey, F. E., and Patterson, A., *Ibid.*, **74**, 4428 (1952); see also ref. 5.
[8] Berg, D., and Patterson, A., *Ibid.*, **75**, (in press) (1953).
[9] Davies, C. W., and James, J. C., *Proc. Roy. Soc.*, **A195**, 116 (1948).

*THE BASIS OF THE OXYGEN EFFECT ON X-IRRADIATED DROSOPHILA SPERM**

BY WILLIAM K. BAKER AND ELIZABETH S. VON HALLE

BIOLOGY DIVISION, OAK RIDGE NATIONAL LABORATORY

Communicated by J. T. Patterson, January 14, 1953

Two hypotheses have been proposed to account for the fact that fewer chromosome aberrations are recovered in organisms exposed to ionizing radiations while in low O_2 concentrations as compared to those treated in air or higher concentrations. On the one hand, the data have been interpreted to mean that there is a differential production in the number of primary breaks with high and low O_2 concentrations.[1-3] On the other hand, the conclusion has been reached that the O_2 is affecting differentially the rejoining process such that more reunion of broken ends occurs in low concentrations.[4] Baker and Edington[5] concluded from studies in Drosophila on the production of recessive lethals and translocations in various O_2 concentrations that the data were compatible with either the differential breakage or the differential reunion hypothesis.

In view of the equivocal results with Drosophila, it seemed advisable to design experiments which would distinguish between the two hypotheses. A study of the induction of dominant lethals could provide the necessary facts. There is strong evidence that the dominant lethals induced in Drosophila sperm are, in the main, the result of inviable chromosome rearrangements (for a discussion of this evidence see Muller[6] and Catcheside[7]). On this basis, the theoretical relationship between dominant lethal production and x-ray dosage has been derived by Lea and Catcheside[8] and extended in a more general form by Haldane and Lea.[9] They conclude that the dose relationship of dominant lethals is very well described by a function with the following three parameters: α, the average number of breaks induced/sperm/1000 r which are available for joining at the time of fertilization; q, the probability that a break will rejoin at fertilization (either restitute or join to form a new arrangement); and D, the dose in kr. Lea and Catcheside[8] found that the theoretical curve which best fitted the experimental data of Catcheside and Lea[10] and Demerec and Fano[11] had the parameters $\alpha = 0.75$ and $q = 0.76$. In these experiments the flies were ex-

posed in air. Now if the differential breakage hypothesis were the valid interpretation of the O_2 effect, then irradiations in N_2 should reduce α but not change q. In view of the reported fact that with chromosome aberrations irradiation in N_2 effectively reduces the dose by a factor of 2.5 as compared to treatment in air, α would be reduced to about 0.3 upon treatment in N_2. In figure 1C is shown the theoretical curves (dotted lines) one would expect to obtain according to the breakage hypothesis when exposures are made in air and N_2.

Material and Methods.—The frequency of dominant lethals, for a given dose of radiation, is subject to such variables as age of males at time of treatment, length of time between irradiation and fertilization, strain used, etc. It is necessary, therefore, to describe in some detail the methods used in the experiments being reported. Virgin Oregon-R males of *Drosophila melanogaster*, which had taken 10–11 days to develop from egg to eclosion, were irradiated 3–4 days after eclosion. Immediately after treatment each male was mated to a single virgin Oregon-R female which was from 4 to 6 days old. Twenty-four hours after mating, the male and female in

TABLE 1

DESIGN OF EXPERIMENTS

DESIGNATION	AGE OF SPERM (HRS.) FROM IRRADIATION TO INSEMINATION	AGE OF SPERM (HRS.) FROM INSEMINATION TO FERTILIZATION	AGE OF SPERM (HRS.) FROM IRRADIATION TO FERTILIZATION
Sperm batch Ia	0–24	0–48	24–48
Sperm batch Ib	0–24	24–72	48–72
Sperm batch IIa	24–48	0–48	48–72
Sperm batch IIb	24–48	24–72	72–96

each vial were separated, the male being mated to a fresh female and the original female being placed alone in a vial containing the egg-laying medium. She was allowed to deposit eggs in this vial for 24 hours (eggs from sperm batch Ia) and then she was transferred to a fresh vial for a final 24 hours of egg laying (sperm batch Ib). The second female with which the male was permitted to mate for 24 hours was also allowed to lay eggs (sperm batch II) over a 2-day period under conditions identical to those of the original female. By use of this design it is possible to follow the dominant lethals induced in at least two successive batches of sperm from the same male as well as to follow the age of the sperm from irradiation to fertilization. In table 1 is shown the manner in which this experimental design relates the egg counts to the age of the sperm.

The egg-laying medium was modified slightly from that described by Carpenter[12] and, in addition, powdered charcoal was added to facilitate counting of the eggs. This medium was poured into 35 × 100-mm. glass vials. A drop of streptomycin solution was added to the surface to cut down bacterial contamination. A 25 × 95-mm. empty glass vial, inverted within

the vial containing the egg-laying medium, served to retain the female and allow her to deposit eggs only within a circular area in the center of the medium. Immediately after the female was removed from the medium, a count was made of the total number of eggs deposited. A count was made of both the hatched and unhatched eggs 24–30 hours later. In some experiments an additional count was made 48 hours after removal of the female but this disclosed no additional hatched eggs (temperature 22–27°C.). Vials in which none of the eggs hatched were not included in the data, since at the moderate dosages used and with the fairly large sample of eggs laid by a female over the 2-day period, practically all these cases must have come from females which were not inseminated.

TABLE 2

EFFECT OF OXYGEN CONCENTRATION AND DOSAGE ON THE INDUCTION OF DOMINANT LETHALS

		SPERM BATCH I		SPERM BATCH II		SPERM BATCH I HELD 24 HOURS	
DOSE (r)	GAS	EGGS COUNTED	PER CENT[a] SURVIVAL	EGGS COUNTED	PER CENT[a] SURVIVAL	EGGS COUNTED	PER CENT[a] SURVIVAL
None	Air	21,336	95.5	20,265	95.6	..	..
1000	N_2	1,477	86.6	1,468	89.5	..	..
	5% O_2	2,027	82.8	1,739	79.7	..	..
	Air	2,250	77.0	2,104	81.3	..	..
2000	N_2	2,895	74.4	3,778	74.4	..	..
	5% O_2	3,812	61.1	4,440	68.4	..	..
	Air	4,232	59.6	3,854	66.9	..	..
4000	N_2	3,763	39.8	3,418	41.7	2,329	43.7
	5% O_2	4,408	24.2	4,037	32.4	2,992	31.4
	Air	4,420	24.1	3,972	30.4	2,501	33.4
7000	N_2	1,504	15.8	1,440	15.4	1,454	18.7
	5% O_2	2,392	6.6	2,229	10.9	1,917	11.1
	Air	1,332	5.1	2,158	11.0	2,246	9.7

[a] The percentage survival given has been corrected for the control survival measured at the same time as the irradiated group.

In each experiment four lots of 20–25 males were used; one lot was not irradiated and served as a control, and the other three lots were irradiated simultaneously in an environment of N_2, 5% O_2 (balance N_2), and air respectively. The various gases were passed over the flies for 10 minutes before, and continuously during, irradiation. Gases and flies were kept at a constant temperature (25 ± 0.5°C.) during treatment. Radiation was administered by a Maxitron machine operating at 250 kvp., 30 ma., 2-mm. Al filter, giving an average intensity of 292 r/min. at the target distance of 49 cm. Dosage measurements were made with a Victoreen 100-r thimble chamber placed in the same location as the treated flies.

Results and Interpretation.—The extensive data gathered in this study have been condensed as much as possible for presentation. Statistical analysis revealed no significant difference between the number of eggs surviv-

ing during the first and second day of egg laying in either the eggs fertilized by sperm batch I or by batch II. Therefore in table 2, which is a compilation of the results, the data from Ia and Ib are lumped as well as the data from IIa and IIb. Also the data from the two or three experiments done at a particular dosage are not kept separate. In the first line of the table are the lumped data from the control experiments, each of which was performed simultaneously with an irradiation. Experiments conducted in N_2, 5% O_2, and pure O_2 under exposure conditions identical to the experimental series except for the lack of radiation gave results similar to the control series in air.

One surprising fact becomes immediately apparent upon examination of the data in this table. The percentage survival is higher in eggs fertilized by sperm of batch II than batch I when the sperm was irradiated in an environment of air or 5% O_2. In N_2, there is no difference in survival of eggs receiving sperm from the two batches. It would thus appear unlikely that the explanation of this phenomenon rests on differential sensitivity, at the time of irradiation, of older and younger sperm. However, as a further check on this point, males were irradiated in the fashion just outlined, but after exposure they were separated individually into vials and held for 24 hours before being allowed to mate. After this period, the experiment was conducted in the same manner as the previous experiments. In the last column of table 1 is given the survival of eggs fertilized with sperm treated in this manner. It is obvious that the survival in this group corresponds to the survival in the second sperm batch and thus an explanation of this phenomenon in terms of differential sensitivity is unequivocally ruled out.

Because of the importance of this increased survival, in interpreting the mechanism of the O_2 effect, it was necessary to determine if this apparent difference rests on a firm statistical basis. Previous workers have shown that the frequency of dominant lethals measured is determined by various extrinsic factors mentioned previously. However, it is not only necessary to hold these factors constant from experiment to experiment, but it is also essential to design the experiments so that variations among the individual males treated can also be taken into account. This is made necessary by the observation that there is much more variation in egg hatch among males given identical treatment than can be accounted for by binomial fluctuations. Therefore, in evaluating the difference between sperm batches, it is necessary to perform an analysis of variance in order that this extraneous variation among males may be taken into consideration. The results of the analysis of variance (arc sine transformation was used) are presented in table 3. It is clearly evident, from the figures in the first line of this table, that eggs fertilized by sperm irradiated in air from batch II definitely have a higher survival value than those fertilized by batch I.

Bonnier and Lüring[13] report just the opposite effect; i.e., a decrease in survival with length of time between irradiation and fertilization. However, an examination of their evidence reveals that, although such a decrease may exist, its effect is not demonstrated until at least 4 or 5 days have elapsed between irradiation and fertilization. They did not observe an increased survival between the first and second day, but this is hardly surprising because of the technique of mass matings used in their experiments. Under such conditions only gross changes would be observable because it is impossible to exclude unhatched eggs laid by virgin females, and because the large extraneous variation among sperm from different males cannot be taken into account.

TABLE 3

ANALYSIS OF VARIANCE OF DATA FROM IRRADIATIONS PERFORMED IN AIR

	1000 r				2000 r			
SOURCE OF VARIATION	DF	MS	F	PROB.	DF	MS	F	PROB.
Sperm I vs. II-treated	1	142.3	4.5	1–5%	1	403.6	9.0	0.1–1%
Sperm I vs. II-controls	1	0.0	0.0	n.s.	1	9.9	0.2	n.s.
Between experiments	1	317.0	11.1	0.1–1%	1	31.7	0.8	n.s.
Controls vs. treated	1	9552.2	334	<0.1%	1	26,468.5	642	<0.1%
Among males[a]	48	28.6			62	41.2		
Males × Sperm I vs. II interaction[b]	48	31.4			62	44.9		

	4000 r				7000 r			
SOURCE OF VARIATION	DF	MS	F	PROB.	DF	MS	F	PROB.
Sperm I vs. II-treated	1	61.6	3.1	5–10%	1	313.3	15.8	<0.1%
Sperm I vs. II-controls	1	19.4	0.5	n.s.	1	177.9	9.0	0.1–1%
Between experiments	2	220.5	3.3	1–5%	1	28.4	1.1	n.s.
Controls vs. treated	1	77,530.4	1172	<0.1%	1	56,774.2	2268	<0.1%
Among males[a]	64	66.1			24	25.0		
Males × Sperm I vs. II interaction[b]	32 C 32 T	38.1[c] 19.9			24	19.8		

[a] These mean squares were used to test significance of variation between experiments and between control vs. treated.

[b] These mean squares were used to test significance of variation between sperm I vs. II in controls and in treated.

[c] Because of the absence of homogeneity among the variances comprising this interaction, the pooled mean squares for controls (C) and treated (T) were used separately in calculating F for sperm I vs. II in controls and in treated, respectively.

The analysis in table 3 also brings out two incidental facts of which we were not previously aware. As can be seen in the second line, there is no

significant difference between sperm batches within the controls except those run with the 7000 r series. This highly significant lowering of the hatch count was brought about by one experiment in which the males and females were allowed to mate for 12 hours instead of the usual 24. This lowered hatch rate was due almost entirely to a decrease in the eggs hatching during the second day of egg laying. Presumably a mating period of 12 hours under the existing conditions was not sufficient to insure that all females would receive sufficient sperm to last for 2 days of laying. The other fact, made evident by the analysis (line three), is that the three experiments at 4000 r are not real replications, as is also true of the two experiments conducted at 1000 r. We are aware of no obvious reason for our inability to replicate experiments in these cases.

One other fact should be noted regarding this measurably time-dependent recovery of sperm from the effect of irradiation conducted in air. As can be seen from table 1, if this recovery was dependent on the total time elapsing from irradiation to fertilization, then sperm batch Ib and IIa should give the same percentage survival of eggs. The data clearly indicate that this is not true, since the survival percentages of eggs fertilized by the same sperm batch but laid over a 2-day period are the same. Therefore, it must be concluded that this recovery is predominantly localized in time to the period from irradiation of the sperm to insemination of the female.

It is impossible, as was pointed out, to explain this recovery in terms of differential sensitivity and thus differential chromosome breakage. However, if one thinks of recovery from breakage in terms of joining of broken chromosome ends, then the results form a consistent picture. It is known that in Drosophila the broken ends do not rejoin to form *new* arrangements until the time of fertilization. Presumably a tandem alignment of the chromosomes in the sperm as observed by Cooper[14] is responsible for this behavior. Cooper concludes (personal communication) that there is not yet sufficient cytological evidence to determine whether this orientation is present in only a few or in most sperm. In any case, such an alignment does not preclude the possibility of restitution of certain breaks. In terms of the differential reunion hypothesis of O_2 action, the dominant lethal data can be interpreted to mean that, although the same number of primary breaks are induced in air and in N_2, the breaks induced in N_2 are more likely to rejoin and to rejoin more quickly than those induced in air. Since the only type of rejoining possible in the sperm is restitution (excluding sister-strand fusion if the chromosome is split) the end-result would be fewer breaks available at the time of fertilization to form new arrangements.

In figure 1 is plotted the observed frequency of surviving eggs as functions of dosage, sperm batch, and the three concentrations of O_2 used. The curves shown in graphs A and C are theoretical and are derived from the formulae given by Haldane and Lea,[9] while those in graph B merely con-

nect the observed points. As mentioned earlier, Lea and Catcheside[8] conclude that their data and those of Demerec and Fano best fit the curve with the parameters $\alpha = 0.75$ and $q = 0.76$. As can be seen in figure 1A, this curve provides a reasonable fit to the data of sperm batch II which come from the type of experiment most comparable to the previous work. In sperm batch I treated in air, it is assumed that not all the potential restitution has taken place, and thus the average number of breaks remaining at the time of fertilization is higher. When $\alpha = 1.0$, the theoretical curve gives a good fit to these empirical points. The N_2 data are in good agreement with a value of $\alpha = 0.62$. It is shown strikingly (Fig. 1C), that the

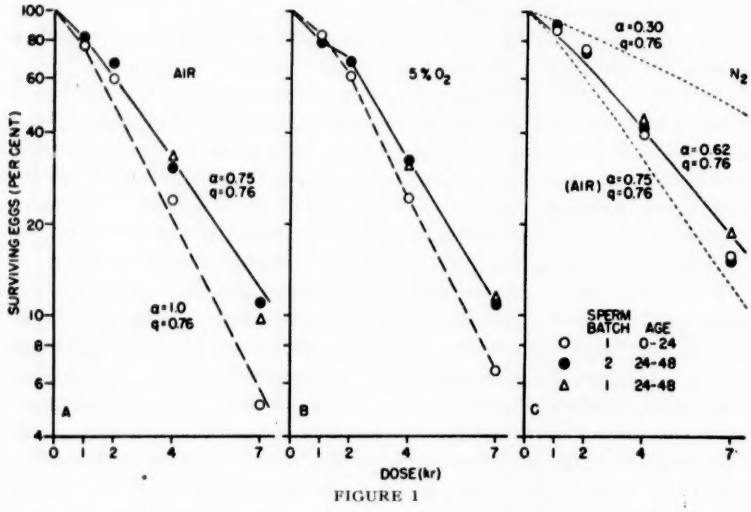


FIGURE 1

Relationship between dosage, oxygen concentration, and sperm batch on the induction of dominant lethals. ○ = sperm batch I; ● = sperm batch II; △ = sperm batch I held 24 hours

experimental points fall far short of the values expected were N_2 causing a two and one-half reduction (as compared to air) in the number of breaks remaining at the time of fertilization.[15] These data, however, give no estimate of the actual number of primary breaks produced (assumed to be independent of O_2 concentration), nor do they allow an estimate of the respective probabilities that a broken end will restitute in the sperm when irradiation is carried out in air or in N_2. From dosage experiments in which insemination follows irradiation very closely, some estimate of the value of these sperm parameters might be possible.

In the foregoing interpretation we have assumed that, although the O_2

concentration affects the ability of a broken chromosome end to restitute in the sperm, it does not influence its ability to rejoin at the time of fertilization. This assumption would appear to be necessary in view of the large O_2 effect (over four times as many translocations recovered in air as in N_2) observed with chromosome translocations in Drosophila.[5] The necessity for this assumption can be seen from the following arguments. A reduction in the value of α from 0.75 to 0.62 is not sufficient, according to the calculations of Lea and Catcheside,[8] to account for a fourfold reduction in the number of viable rearrangements. These calculations also show that if the same number of breaks were available at the time of fertilization in N_2-treated and air-treated sperm and if in the former rejoining of broken ends was more likely, then more rearrangements should be recovered in the N_2-treated sperm than the ones exposed in air. Therefore, if the O_2 concentration also affected the rejoining ability at fertilization, a still smaller difference would be expected between the frequency of translocations recovered in air and in N_2. There still remains the problem as to the cause of the large O_2 effect with translocations. It is possible that this discrepancy between the dominant lethal and the translocation data may be caused by the fact that the latter experiments were conducted at 8°C., the former at 25°. Previous work[16] with recessive lethals indicated a larger O_2 effect at the low temperature, but whether or not this is the explanation will have to await further experimentation.

Discussion.—Riley, Giles, and Beatty[3] argue against the reunion hypothesis of O_2 action on the basis of two lines of evidence: (1) Experiments conducted using Tradescantia (Giles and Riley[1]) have shown that changing the O_2 concentration immediately after irradiation, when some of the induced breaks still remain free, has no effect on the yield of chromosome aberrations. (2) The yield of chromatid exchanges varies with intensity of radiation in a similar manner when the irradiation is carried out in N_2 or in O_2.[3] Giles and Riley point out, concerning the first line of evidence, that the experiments do not exclude the possibility that the subsequent behavior of a broken end in rejoining is determined by conditions existing at the time the break is produced. Thus under this possibility, only the O_2 concentration at the time of irradiation would affect the yield of aberrations. The second line of evidence is taken by Riley, *et al.*,[3] to indicate that, "the average restitution time is essentially the same for breaks produced in the presence or in the absence of oxygen." However, strictly speaking, the only conclusion that can be implied from their experiment is that the breaks which are to be used *in exchanges* remain open for approximately the same length of time in O_2 and in N_2. Thus the evidence has no direct bearing on the question as to whether breaks formed in N_2 are more likely to restitute than those induced in O_2 or whether the restitution takes place more quickly. Also it should be noted that while the frequency of chromatid ex-

changes increases with intensity, there is no compensatory decrease in the frequency of chromatid breaks as might be expected if, as is generally assumed, the exchanges arise from two originally produced chromatid breaks. This may mean that there are two types of breaks produced: those which can either restitute or join in a new arrangement, and those which remain permanently broken and subsequently appear as chromatid or isochromatid breaks.

It is generally thought that chromatid deletions and exchanges come from chromatid breaks. However, consideration should be given to the likely possibility that isochromatid breaks can also produce deletions and exchanges when one or more of the broken chromatids restitute. If we keep this possibility in mind, and if we assume that the lack of O_2 makes more likely the restitution of breaks capable of joining, then certain observations of Riley, *et al.*,[3] become more meaningful. A lowering of the O_2 tension during irradiation would markedly decrease the frequency of chromatid exchanges and isochromatid deletions. However, it would be expected that the frequency of chromatid deletions would not be lowered as much since restitution of one of the broken chromatids in an originally produced isochromatid break would increase the number of recovered chromatid deletions. These workers observed a much lower O_2-He dose ratio in the case of chromatid deletions as compared with either exchanges or isochromatid breaks. As they point out, such a difference would not be expected on the basis that O_2 concentration is affecting the initial number of breaks induced.

Support for the reunion hypothesis comes from the following four observations: (1) The very low frequency of endosperm mosaics formed (in contrast with interstitial deletions) in maize when irradiation is carried out in N_2 as compared to air (Schwartz,[4] and unpublished). These mosaics are formed when breaks do not restitute and sister-strand fusion initiates the chromatid type of bridge-breakage-fusion cycle. (2) Dominant lethals in Drosophila are not reduced with irradiation in N_2 to nearly the extent one would expect on the breakage hypothesis. (3) The demonstrated recovery from dominant lethals (interpreted as restitution of some breaks) which is dependent upon the time elasping between irradiation and insemination in sperm treated in air but not demonstrably dependent in N_2-treated sperm. (4) The previously mentioned low dose reduction observed with chromatid deletions in Tradescantia upon exposure in He.

From the available data, there is little reason to believe that the frequency of induced primary breaks is affected by O_2 concentration. Rather, we envisage the substances induced by the interaction of O_2 and x-rays as ones which produce chromosome breaks less capable of subsequent rejoining than the breaks induced in the absence of O_2. In Drosophila, these substances are effective only on the process of restitution, and apparently extend the average length of time between breakage and restitution. It re-

mains to be seen whether the O_2 concentration affects the general rejoining in other organisms in which the process of restitution and rejoining to form new arrangements are not separated in time.

Summary.—A study of the relation between x-ray dosage and O_2 concentration on the induction of dominant lethals in mature sperm of *D. melanogaster* shows that the frequency of dominant lethals induced upon irradiation in N_2 is reduced much less than would be expected on the hypothesis that O_2 concentration is affecting the number of primary breaks induced. Fewer dominant lethals are recovered in sperm exposed in air when at least 24 hours have elapsed between treatment and insemination than when insemination very shortly follows treatment. This effect is not observed in N_2-treated sperm. These data can be interpreted on the basis that the O_2 concentration affects the amount of restitution of chromosome breaks taking place in the sperm. A low O_2 concentration during irradiation makes restitution more likely, and the broken ends apparently restitute more quickly. Therefore, the data lend support to the differential reunion hypothesis of O_2 action rather than to the differential breakage hypothesis.

Acknowledgment.—The authors are indebted to Dr. A. W. Kimball and Mr. G. J. Atta for conducting the tedious statistical analyses of the data.

* Work performed under Contract No. W-7405-eng-26 for the Atomic Energy Commission.

[1] Giles, N. H., and Riley, H. P., these PROCEEDINGS, **36,** 337–344 (1950).

[2] Hollaender, A., Baker, W. K., and Anderson, E. H., *Cold Spring Harbor Symp. Quant. Biol.*, **16,** 315–325 (1951).

[3] Riley, H. P., Giles, N. H., and Beatty, A. V., *Am. J. Botany*, **39,** 592–596 (1952).

[4] Schwartz, D., these PROCEEDINGS, **38,** 490–494 (1952).

[5] Baker, W. K., and Edington, C. W., *Genetics*, **37,** 665–677 (1952).

[6] Muller, H. J., *J. Genetics*, **40,** 1–66 (1940).

[7] Catcheside, D. G., *Adv. Genetics*, **2,** 271–358 (1948).

[8] Lea, D. E., and Catcheside, D. G., *J. Genetics*, **47,** 10–24 (1945).

[9] Haldane, J. B. S., and Lea, D. E., *Ibid.*, **48,** 1–10 (1947).

[10] Catcheside, D. G., and Lea, D. E., *Ibid.*, **47,** 1–9 (1945).

[11] Demerec, M., and Fano, U., *Genetics*, **29,** 348–360 (1944).

[12] Carpenter, J. M., *Dros. Info. Service*, **24,** 96–97 (1950).

[13] Bonnier, G., and Lüning, K. G., *Hereditas*, **36,** 445–456 (1950).

[14] Cooper, K. W., *Yearbook Am. Phil. Soc. for* **1951**, pp. 146–147.

[15] This small oxygen effect is in line with the observation of Russell, *et al.* (Russell, W. L., Kile, J. C., and Russell, L. B., *Genetics*, **36,** 574, 1951) on the induction of dominant lethals in mice. They observed no difference in the frequency of dominant lethals induced in air and in 2% O_2.

[16] Baker, W. K., and Sgourakis, E., these PROCEEDINGS, **36,** 176–184 (1950).

THE GENETICS OF HOMEOSTASIS IN DROSOPHILA*

BY TH. DOBZHANSKY AND BRUCE WALLACE

COLUMBIA UNIVERSITY, NEW YORK CITY, AND BIOLOGICAL LABORATORY, COLD SPRING HARBOR, N. Y.

Communicated January 13, 1953

The contribution of an individual of one generation to the gene pool of the next depends upon the ability of the individual to survive and to reproduce. This contribution, relative to that of other individuals of the same population, is a measure of the biological adaptive value of the individual in the succession of environments which it has encountered up to the close of its reproductive age. Similarly, the adaptive value of a population is determined by the average survival and reproductive capacities of the many gene combinations which are produced in this population. Since the environment is everchanging, an organism is exposed to a greater or lesser variety of environments. An organism is usually capable of carrying on its vital functions in all environments in which the species or population to which it belongs normally exists. Adaptation to a variety of environments is accomplished in two ways. First, most species and populations are polymorphic and consist of a variety of genotypes optimally adapted to different aspects and sequences of environments. Secondly, individuals respond to environmental changes by physiological and structural modifications. Modifications evoked by environmental variations recurrent in the environment of the species almost always tend to increase the probability of survival and reproduction of the organism. The organism adjusts itself to recurrent environmental changes in such a way that its functioning continues unimpaired; it is said to be homeostatic.[1]

Homeostatic mechanisms have been carefully studied by physiologists, especially in the higher vertebrates. It is nevertheless far from universally realized that homeostasis is conditioned by the genotype, and the different genotypes permit different degrees of homeostasis. The "wisdom of the body" is an outcome of the molding of the genetic structure of the species by natural selection in the process of evolution, and it cannot be understood outside this evolutionary context. As pointed out by several authors, especially by Schmalhausen,[2] genotypes which are favored most strongly by natural selection are those which condition homeostatic responses to recurrent environmental stimuli. Stimuli which were seldom or never encountered in the history of the species very often evoke adaptivity indifferent or even positively harmful responses, which Schmalhausen has called morphoses. Mutants and gene combinations which have not been historically established as normal constituents of natural populations of the species may be deficient in homeostatic responses even under usual

environments. These new or rare genotypes have not yet become fitted by natural selection to any particular environment. Homeostasis, like organic adaptedness in general, is not an inherent attribute of living matter as is often assumed by vitalists and by Lamarkians of various kinds.

The genetic mechanisms which underlie homeostasis have been explored very little. For obvious technical reasons, classical genetics preferred to deal with clean-cut genetic differentials regardless of their adaptive significance. We have been led to this problem by studies on the genetics of natural and experimental populations. Here the homeostatic properties of some genotypes contrast with a relative lack of such properties in others.

Experimental Procedures.—Four species of Drosophila are involved in

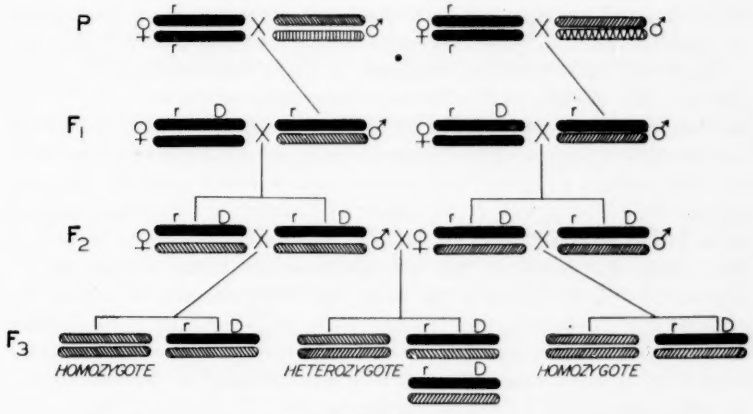


FIGURE 1

A generalized scheme of crosses used to obtain homozygotes and heterozygotes for various autosomes from wild and experimental populations. Black: chromosomes with mutant genes used as markers; *r*: recessive mutant; D: dominant mutant. Hachured: chromosomes, the effects of which are analyzed.

the present investigation, namely *D. pseudoobscura*, *D. persimilis*, *D. prosaltans*, and *D. melanogaster*. Samples of the natural populations of the first and second species were collected in the summer of 1951 in the Yosemite National Park region of California. Samples of *D. prosaltans* came from several localities in different parts of Brazil (Pirassununga, state of São Paulo, collected by Drs. C. Pavan and A. B. da Cunha; Ferreira Gomes, territory of Amapá, northeastern portion of the Isle of Marajó, state of Pará; Fordlandia, Pará; Içana, upper Rio Negro, Amazonas; and upper Rio Doce, Minas Geraes, collections of Dr. A. B. da Cunha and Th. Dobzhansky, made in 1952). All these population samples were brought or sent to the laboratory in New York and used to study the con-

cealed genetic variability due to autosmal recessive genes and gene combinations carried in the populations in question. *D. pseudoobscura* and *D. persimilis* have three pairs of large autosomes (the second, third, and and fourth chromosomes), and a pair of microchromosomes (the fifth chromosome). *D. prosaltans* has two pairs of large autosomes (the second and the third). The genetic variability in all of these autosomes, except the microsomes, has been studied. The results obtained will be described in detail elsewhere. Here it will be sufficient to state that in all cases a certain proportion of the chromosomes proved to be lethal in double dose, i.e., in homozygotes. Other chromosomes were semilethal, or subvital, or caused sterility, or modifications of the developmental rates, or morphological aberrations in the homozygotes. More important at present is that the techniques used in these studies were alike in principle in all cases, the essential features of these techniques being as follows (Fig. 1):

Males collected in natural habitats, or single sons of the females so collected, are crossed to females of a laboratory strain homozygous for suitable recessive mutant genes (*r*, in Fig. 1). A single male from each progeny (F_1) is outcrossed to females carrying in one of their chromosomes the same recessive marker, a dominant marker which is lethal to homozygotes (*D*, in Fig. 1) and an inversion which suppresses the recombination in the proper chromosome. In the next generation (F_2) flies of both sexes which show the dominant but not the recessive marker are selected in each strain. Such flies carry the same "wild" chromosome, as well as the chromosome with the marking genes. When they are inbred, one-third of their progeny (F_3) is expected to be homozygous for a "wild" chromosome (i. e., to carry two replicas of a certain chromosome descended from the wild ancestor). Two thirds of the progeny should carry this "wild" chromosome and the chromosome with the marker genes (Fig. 1). In reality, the homozygotes are often less numerous than expected because some of the "wild" chromosomes are deleterious to the homozygotes. If no homozygotes appear, the chromosome is lethal in double dose; if less than half of the expected proportion of the homozygotes survive, the chromosome is semilethal. (For more details about the method see previous publications.[3–6] For our present purposes the lethal and semilethal chromosomes may be disregarded. The remaining chromosomes, very roughly three-quarters of the total, fall within an approximately bell-shaped curve, somewhat truncated on the left owing to the exclusion of semilethals. It is with these remaining chromosomes, which may be conditionally referred to as "quasi-normal" that we are at present concerned.

The different quasi-normal chromosomes give different proportions of the homozygotes in the test cultures. Apart from the variability due to sampling errors, these differences are the result of two causes. First, the homozygotes for some of these chromosomes are more viable than others;

with sufficient data one can distinguish chromosomes which are subvital, normal, and supervital when homozygous. Secondly, the environment, the culture conditions, in which the tests are made are not uniform. Thus, some cultures are more crowded than others, or have unequal amounts of food, etc. The importance of this environmental component of the variance can be estimated from our data. Indeed, the tests were so arranged that the cultures of the final generation (the lowermost line in Fig. 1) were replicated for most of the chromosomes tested. A group of parents (F_2) were permitted to oviposit in a culture for 2–3 days, whereupon the parents were transferred for oviposition in a new culture bottle. For the different species and chromosomes, from one to five such transfers were made, giving from two to six replicate cultures. Now, if the viability of the homozygotes for a given chromosome is not sensitive to the environmental variations encountered in the replicate cultures, these cultures will give the same proportions of the homozygotes, within limits of the sampling errors. Conversely, an environmental sensitivity may result in statistically significant heterogeneities between the replicate cultures. This can be detected by the Brandt-Snedecor chi-square test for homogeneity among the replications.

Homozygosis for any chromosome is rare in any population of most species of Drosophila. In all but the very closely inbred populations the two chromosomes of each pair carried by most individuals differ in origin and thus have somewhat different gene complexes. Wild flies are usually heterozygotes rather than homozygotes. Corresponding heterozygotes can be obtained also in laboratory experiments. In the F_2 generation of the crosses shown in figure 1 one must now take females and males from the offspring of *different* wild progenitors. Such experiments have actually been made, using chromosomes which were quasi-normal as well as those which were lethal or semilethal when homozygous (see the middle of the F_2 and F_3 of Fig. 1). In these experiments the proportions of the flies heterozygous for the two "wild" chromosomes formed a nearly symmetrical bell-shaped distribution. As with homozygotes, the variance observed in this distribution may be the result of three causes. Apart from the errors of sampling, different chromosome combinations may give flies of different viability under one environment, or similar combinations may give different viabilities under different environmental conditions. The experiments with chromosomal heterozygotes were arranged like those that involved homozygotes. Most chromosome combinations were obtained in replicate cultures. The sensitivity of the heterozygotes to different environmental conditions may be tested by computing chi-squares for homogeneity of the replications.

The experiments involving *D. melanogaster* were somewhat differently arranged.[6, 7] During the course of regular sample analyses—similar in

technique and purpose to those described above for natural populations—an additional effort was made to find combinations of second chromosomes that exhibited "negative heterosis." The original matings that yielded homozygous and heterozygous wild type individuals were set up in such a way that the viability of individuals heterozygous for a given pair of second chromosomes could be compared with the viability of individuals homozygous for each of the two members of the pair. These matings were not replicated.

TABLE 1

SENSITIVITY TO ENVIRONMENTAL VARIATIONS SHOWN BY HOMOZYGOTES AND HETEROZYGOTES FOR CHROMOSOMES DERIVED FROM CERTAIN NATURAL POPULATIONS

		HOMOZYGOTES			HETEROZYGOTES		
SPECIES	CHROMOSOME	CHI-SQUARE	DEGREES OF FREEDOM	P	CHI-SQUARE	DEGREES OF FREEDOM	P
Pseudoobscura	II	144.34	81	<0.001	31.91	33	0.5
Pseudoobscura	III	212.24	86	<0.001	33.29	31	0.3
Pseudoobscura	IV	173.86	77	<0.001	44.09	30	0.05
Persimilis	II	225.77	184	0.018	74.92	99	0.96
Persimilis	III	519.28	329	<0.001	67.27	63	0.3
Persimilis	IV	327.43	269	0.008	99.05	66	0.004
Prosaltans	II	253.14	152	<0.001	48.07	47	0.4
Prosaltans	III	290.40	216	<0.001	72.36	54	0.046

TABLE 2

SENSITIVITY TO ENVIRONMENTAL VARIATIONS SHOWN BY HOMOZYGOTES AND HETEROZYGOTES FOR SECOND CHROMOSOMES FROM EXPERIMENTAL POPULATIONS OF *Drosophila melanogaster*. N IS THE NUMBER OF THE CHROMOSOMES OR CHROMOSOME COMBINATIONS STUDIED

	HOMOZYGOTES				HETEROZYGOTES			
POPULATION	N	CHI-SQUARE	DEGREES OF FREEDOM	P	N	CHI-SQUARE	DEGREES OF FREEDOM	P
1	20	88.35	75	0.50	11	32.16	40	>0.50
3	26	141.38	102	0.01	13	53.81	51	0.50
5	7	32.47	23	0.09	4	17.61	13	0.18
6	4	18.51	16	0.30	2	14.80	8	0.07
7	40	168.39	139	0.09	23	71.20	82	>0.50
TOTAL	97	449.10	355	<0.001	53	189.58	194	>0.50

Whenever a combination of cultures was found, however, in which the viability of the heterozygous individuals appeared to be less than that of the two homozygotes, replicate F_4 cultures were made using $D/+F_3$ (*Cy L*/+ in the case of *D. melanogaster*) flies as parents. Usually four or five, rarely three, replicate cultures of each type of homozygous and of the heterozygous crosses were made. The final decision as to the presence or absence of negative heterosis was based on these greatly increased numbers of flies. This slightly unorthodox reason for deciding which chromosomes to

test in replicate cultures would have the effect of selecting those chromosomes which when homozygous resulted in higher than average viabilities and chromosomal combinations that produced correspondingly lower viabilities. The Brandt-Snedecor chi-square test for homogeneity was used among these replications as it was in those described above.

Environmental Sensitivity of Homozygotes and Heterozygotes.—A summary of the results of the homogeneity tests for the replicate cultures is presented in tables 1 and 2. In both tables the cultures producing flies homozygous for certain chromosomes derived from natural or experimental irradiated populations are contrasted with cultures giving rise to flies with the two chromosomes of a pair derived from different progenitors.

It can be seen at a glance that the homozygotes generally show very significant heterogeneous viabilities in replicate cultures, while the heterozygotes usually do not. This situation has been emphasized in table 2 by summing the chi-squares and degrees of freedom for each type of replication. The individual entries in table 1, with few exceptions, agree in demonstrating the heterogeneity among homozygous replications and the homogeneity among the sets of heterozygotes. The data in table 2 are not so consistent; it is quite probable that replicated cultures of chromosomes from population 1, for instance, are not heterogeneous. The difference between tables 1 and 2 may result from any of three factors: (1) The number of *D. melanogaster* chromosomes analyzed was quite small. (2) The choice of chromosomes tested in replicate cultures from the experimental populations favored homozygotes with high viabilities. (3) The experimental irradiated populations of *D. melanogaster* were kept under relatively constant conditions where selection for homeostasis would be minimized; indeed, the Oregon-R strain of flies from which these populations were started has been kept under laboratory conditions for at least 25 years.

Environmental Sensitivity of Different Chromosomes.—Dobzhansky and Spassky[4] measured the viability of homozygotes for 26 second and 22 fourth chromosomes of *D. pseudoobscura* at three different temperatures: $16^1/_2°$, 21°, and $25^1/_2°$. They found that about half of these chromosomes showed significant differences in performance at the three temperatures, while the other half were not temperature sensitive. Their experiments, like those described in the present article, involved raising replicate cultures for every chromosome. Some of the chromosomes showed quite significant heterogeneities between the replicate cultures at all temperatures; others were heterogeneous at only one temperature; still others gave no heterogeneities. The environmental variable that caused the heterogeneity was in some instances quite obvious: different degrees of crowding in replicate cultures. The homozygotes for some chromosomes are sensitive to crowding and survive rather poorly in crowded cultures; other chromo-

somes seem to be insensitive to crowding, at least within the limits studied. When the heterogeneity is not correlated with crowding, it must be due to more subtle differences in culture conditions. Nutritional variables, such as different microflorae in different cultures, may be involved.

Examination of our data discloses a situation similar to that observed by the authors just cited. The contributions of some of the chromosomes to the chi-squares are much greater than those of other chromosomes. The over-all evidence shows that, on the average, the homozygotes possess less perfect homeostatic properties than do the heterozygotes; the evidence is not critical, however, in determining whether all homozygotes are deficient in homeostasis or whether some homozygotes are as adaptable as most heterozygotes (see, for instance, population 1, table 2). We have by no means explored the norms of reaction of the genotypes in question. It should be noted that in these experiments the environment was deliberately made as uniform as practicable in fairly large-scale work. Nevertheless, many homozygous genotypes proved to be sensitive to residual environmental diversities that failed to alter the viabilities of heterozygous individuals.

Discussion.—Cannon,[1] the greatest student of homeostasis, wrote as follows: "In an open system, such as our bodies represent, compounded of unstable materials and subjected continually to disturbing conditions, constancy is in itself evidence that agencies are acting, or ready to act, to maintain this constancy." And: "It is not supposed that the full display of homeostatic adjustments will be found in all forms of animals." Homeostasis is a matter of degree. In general, homeostasis with its consequent autonomy of life processes must confer a high selective advantage in most species. We do not know which physiological processes in Drosophila must be maintained constant to enable the development to proceed unimpaired, nor do we know the nature of the gene action that produces the buffered system for the constancy of such processes. The evidence of homeostasis is necessarily indirect. Nevertheless, heterozygotes are more uniformly successful in a variety of environments than are homozygotes; this suggests that the heterozygotes are better able than homozygotes to cope with these different environments and to maintain their internal milieu in functional order. Wigan[8] and, more recently, Robertson and Reeve[9] have reported similar phenomena within *D. melanogaster* and Mather[18] in *Primula sinensis*. In Robertson and Reeve's report it has been shown that environmentally caused variability for certain morphological characteristics is inversely proportional to the degree of heterozygosity. The generality of this correlation between homeostasis and heterozygosity is, of course, a matter of speculation at present. It should be emphasized, however, that our data are based on the relative frequencies of two classes of flies ($D/+$ and $+/+$) and, consequently, statistically homogeneous data

within sets of replications of heterozygous combinations indicate the existence of homeostasis within *both* classes of flies. The fact that a pronounced homeostasis exists among heterozygous individuals helps in our understanding of the genetic structure of populations. Within recent years it has been recognized that the genotype of a Mendelian population is a coadapted system. The evidence of this coadaptation has been obtained from studies on heterosis in inversion heterozygotes,[10, 11] on the adaptive values of experimental populations,[6, 7, 12] and on the breakdown of heterosis in the F_2 generation hybrids between natural populations of different localities.[13] Additional evidence[10, 12] has indicated that this coadaptation is based upon a genetic heterogeneity—not upon genetic unifomity and homozygosis within a population.

If heterozygosity—at least within some as yet unknown limits—results in an increase of homeostasis, then the extreme genetic heterogeneity within a population can be understood: selection for multiple alleles at many loci would act to minimize the frequency of homozygosity at any one locus. Coadaptation, then, leads to the formation of a gene pool containing those members of each series of alleles that are most likely to produce harmonious combinations with other alleles at the same locus and in combination with the alleles at all other loci. Selection must operate to choose a variety of alleles at every locus, thus avoiding the extreme sensitivity to environmental differences which characterizes many homozygotes; but it must also weed out those extreme variant alleles which often react with others at the same or other loci to produce morphoses. This results in the accumulation of a store of mutually compatible alleles—in other words, in a coadapted genetic system.

Note should be taken of the fact that not all of our data on heterozygotes show evidence of complete homeostasis. A significant chi-square (table 1) was found for fourth-chromosome heterozygotes in *D. persimilis*, and chi-squares that lie at the boundary of the conventional significance range for the fourth-chromosome heterozygotes in *D. pseudoobscura* and the third chromosome in *D. prosaltans*. This is not unexpected. First of all as mentioned above, in our experiments the viability of the homo- and heterozygotes for the wild chromosomes is measured against that of carriers of certain dominant mutant genes introduced as markers (*D*, in Fig. 1) and some of the statistical heterogeneity may arise from a lack of homeostasis in the mutant heterozygotes. Second, all heterozygotes need not possess equal homeostatic capacities; the conditions within standard laboratory cultures may well represent an environment that falls outside the range of those normally encountered in natural habitats. Furthermore, the coadaptation of the gene complexes in a population need not be so perfect that all individuals that carry two chromosomes from the same population will be equally versatile. A heterogeneity of viabilities of

heterozygotes has, indeed, been observed by Cordeiro[14] and by Wallace and King.[7] Lastly, occasionally inbreeding or the operation of chance must produce some individuals in nature—and in our heterozygous combinations as well—that carry a chromosome section in duplicate and are homozygous for that section.

Natural selection is to a great extent opportunistic, although interpopulation competition serves to limit this opportunism somewhat. Nevertheless, it is impossible to postulate abstract schemes of evolution without regard to the biology of the group in which selection is acting. It may be expected that coadapted gene pools and coadaptation of random combinations of chromosomes arising from these pools would be characteristic of crossfertilizing species with rather large effective population sizes. If inbreeding or selfing is the rule, it may be expected that homozygotes would show consistently high adaptive values in a variety of environments (see, for instance, Gustafsson[15]). Higher animals with their low fecundity and individual longevity may be expected to develop a different integration of their gene pools than would perennial plants with their great numbers of seeds and alternative asexual methods of perpetuation. Finally, microorganisms with their enormous reproductive potentials and extremely short generation time may rely upon a still different scheme; here gene mutation may serve as the main adaptive mechanism.

Summary.—We have studied the rates of survival in crowded cultures of individuals "homozygous" and "heterozygous" for chromosomes derived from natural and experimental populations of *Drosophila pseudoobscura, D. persimilis, D. prosaltans,* and *D. melanogaster.* The "homozygotes" carried certain chromosomes in duplicate, while the "heterozygotes" had the two chromosomes of a pair derived from different wild progenitors, or from different members of an experimental population. The experiments were so arranged that each chromosome combination was tested in replicate cultures. The environments in the replicate cultures varied because of different degrees of crowding and probably also because of variations in the quantity and quality of the food. The homozygotes often showed significantly different survival rates in the varying environments of the replicate cultures. Conversely, the heterozygotes gave usually uniform survival despite the environmental variations. It is inferred that the homeostatic adjustments are superior in heterozygotes than in homozygotes. The gene complexes carried in the homologous chromosomes in sexual and crossfertilizing populations are coadapted by natural selection to give high fitness in heterozygotes with most other chromosomes of the same population. The genotype of a Mendelian population is an integrated system, the parts of which are fitted together in the process of evolution.

Acknowledgments.—We are indebted to Mrs. O. A. Pavlovsky, N. P.

Dobzhansky, Carol Madden, and to Mr. B. A. Spassky for their efficient help in carrying out the very laborious experiments and calculations. Professor Howard Levene and Mr. Timothy Prout have contributed some valuable advice concerning statistical problems. We wish also to acknowledge stimulating discussions with Professor I. Michael Lerner, who has dealt with the problem of "genetic homeostasis" in previous publications.[16,17]

* The work reported in this article has been carried out under Contract No. AT-(30-1)-1151 and AT-(30-1)-557, U.S. Atomic Energy Commission. The collection of the population samples of *Drosophila prosaltans* in Brazil has been made possible thanks to a grant from the Rockefeller Foundation.

[1] Cannon, W. B., *The Wisdom of the Body*, W. W. Norton, New York, 1952.
[2] Schmalhausen, I. I., *Factors of Evolution*, Blakiston, Philadelphia, 1949.
[3] Dobzhansky, Th., Holz, A. M., and Spassky, B., *Genetics*, **27**, 463–490 (1942).
[4] Dobzhansky, Th., and Spassky, B., *Ibid.*, **29**, 270–290 (1944).
[5] Pavan, C., *et al.*, *Ibid.*, **36**, 13–30 (1951).
[6] Wallace, B., and King, J. C., *Am. Nat.*, **85**, 209–222 (1951).
[7] Wallace, B., and King, J. C., PROC. NATL. ACAD. SCI., **38**, 706–715 (1952).
[8] Wigan, L. G., *J. Genetics*, **46**, 150–160 (1944).
[9] Robertson, F. W., and Reeve, E., *Nature*, **170**, 286 (1952).
[10] Dobzhansky, Th., *Genetics*, **35**, 288–302 (1950).
[11] Dobzhansky, Th., and Levene, H., *Am. Nat.*, **85**, 247–264 (1951).
[12] Wallace, B., et al., *Genetics* (in press).
[13] Vetukhov, M. A., PROC. NATL. ACAD. SCI., **39**, 30–34(1953).
[14] Cordeiro, A. R., *Ibid.*, **38**, 471–478 (1952).
[15] Gustafsson, A., *Hereditas*, **32**, 263–286 (1946).
[16] Lerner, I. M., *Population Genetics and Animal Improvement*, Cambridge University Press, London, 1950.
[17] Lerner, I. M., and Demster, E. R., *Heredity*, **5**, 75–94 (1951).
[18] Mather, K., *Evolution*, **4**, 340–352 (1950).

---

*HETEROGENEITY OF CLONES OF SACCHAROMYCES DERIVED FROM HAPLOID ASCOSPORES**

BY HERSCHEL ROMAN AND STANLEY M. SANDS

DEPARTMENT OF BOTANY, UNIVERSITY OF WASHINGTON, SEATTLE

Communicated by L. J. Stadler, January 21, 1953

It has been demonstrated[1, 2] in heterothallic strains of Saccharomyces, that crosses between clones derived from haploid ascospores sometimes produce asci which exhibit tetraploid segregation. One explanation for the occurrence of such asci is that they arise from tetraploid ascogenous cells which are in turn the result of fusion of diploid cells present in the parental clones.[1] This possibility points up the need for information regarding the composition of parental clones of haploid origin, in order to achieve

a better understanding of segregations obtained from crosses. As a start in this direction, we have investigated a number of such clones for evidence of diploidy. The results are given below.

Material.—We are indebted to Dr. C. C. Lindegren for providing us with clones 1 and 21, which are ancestral to the other clones examined for heterogeneity. Clone 1 is a "rapid" galactose fermenter[3] and exhibits the "flaky" character[4] in liquid medium; both of these characters are dominant. Clone 21 is a slow fermenter and is non-flaky. When the two clones were first received, clone 1 was of mating-type **α** and clone 21 of mating-type **a**. Crosses between the two strains gave, with rare exceptions, asci

TABLE 1

THE TYPES OF SECONDARY CLONES OBTAINED FROM LARGE AND SMALL CELLS ISOLATED FROM CLONES OF SACCHAROMYCES GROWN FROM HAPLOID ASCOSPORES

PARENTAL CLONE	SMALL-CELLED SECONDARY CLONES DERIVED FROM		LARGE-CELLED SECONDARY CLONES DERIVED FROM	
	LARGE CELLS	SMALL CELLS	LARGE CELLS	SMALL CELLS
1	0	0	6	6
21	0	0	5	8
4	3	7	7	0
27	5	14	10	0
260	0	8	12	1

TABLE 2

RESULTS OF MATING-TYPE TESTS OF SECONDARY CLONES DESCRIBED IN TEXT

PARENTAL CLONE	ORIGINAL GENOTYPE	TYPE OF SECONDARY CLONE	MATING TYPE OF SECONDARY CLONES		
			a	α	NON-MATER
4	*αgF*	Small-celled	0	10	0
27	*aGf*	Small-celled	14	0	0
260	*aGf*	Small-celled	8	0	0
4	*αgF*	Large-celled	0	7	0
27	*aGf*	Large-celled	0	0	10
260	*aGf*	Large-celled	3	0	10
1	*αGF*	Large-celled	0	6	6
21	*agf*	Large-celled	0	0	13

with 2:2 segregations for mating type, galactose fermentation, and flaky. Thus clones 1 and 21 were originally both haploid and their original genotypes may be represented as *αGF* and *agf*, respectively.

Evidence of Genetic Heterogeneity.—In this material, a clone grown from a haploid ascospore at first consists of small, spherical cells which average about 4 microns in diameter and which grow in clusters (the "short-shoot" growth habit described by Winge and Laustsen[5]). Numerous crosses between clones at this stage have given segregations which indicate that the cells are haploid. As the clone continues to grow, cells of larger size, about 7 microns in diameter, make their appearance and an older clone, especially after several subcultures, often contains a preponderance of such cells.

Strains which have become heterogeneous in this fashion exhibit a diminished capacity to produce zygotes in crosses with strains of opposite mating type. For example, clone 21 lost this capacity almost completely after a series of subcultures, although it was at first a vigorous zygote-producer in crosses.

Five clones which were derived from haploid ascospores and which had become heterogeneous for cell size were selected in our first attempts to obtain evidence of diploidy. Both large and small cells were isolated with a micromanipulator and placed on agar droplets.[6] After 1–2 days growth on the droplets, the colonies (hereafter called secondary clones) were transferred to slants and examined two days later for cell size. In 29 of 30 secondary clones obtained from small cells of clones 4, 27, and 260, the cell size and growth habit were of the type associated above with the haploid condition (table 1). On the other hand, 29 of the 37 secondary clones

TABLE 3

THE COMPOSITION OF VARIOUS PARENTAL CLONES SAMPLED BY THE PLATING METHOD

PARENTAL CLONE	ORIGINAL GENOTYPE	TYPES OF SECONDARY CLONES: SMALL-CELLED a	SMALL-CELLED α	SMALL-CELLED NON-MATER	LARGE-CELLED a	LARGE-CELLED α	LARGE-CELLED NON-MATER
242A	*aGF*	44	0	0	0	0	0
B	*aGF*	40	0	0	0	0	0
C	*αGF*	0	40	0	0	0	0
D	*αGF*	0	25	0	0	0	0
1	*αGF*	0	7	0	0	21	86
13	*αgf*	0	8	0	0	3	25
15	*αGf*	0	25	0	0	0	11
21	*agf*	0	0	0	2	0	62
27	*aGf*	4	0	0	0	0	44
30	*agF*	6	0	0	0	0	30
260	*aGf*	2	0	0	1	0	45

grown from large cells were predominantly large-celled themselves. In these there was actually a range in cell size and some of the cells were as small as those of haploid clones. It will be seen later, however, that the range in cell size is characteristic of diploid clones and is not due to the presence of haploid cells. Clones 1 and 21, which were also selected for examination, gave only large-celled secondary clones, even when these were derived from small cells. As table 3 shows, this may be attributed to the fact that clones 1 and 21 had become diploid to such an extent that the small cells isolated from these clones were actually diploid cells that were either arrested in growth or had not yet attained full size at the time of selection.

The secondary clones were tested next for mating type (table 2). All of the small-celled clones exhibited a mating reaction and in each case the mating type was that of the parental clone from which the secondary clone was derived. The large-celled clones, however, were of two types: one

gave the expected mating reaction whereas the other failed to mate with either **a** or **α** mating-type strains. As a safeguard against the possibility of contamination, all of the secondary clones were tested also with respect to galactose fermentation and the flaky character and all gave results identical with those of the parental clones from which they were isolated.

Aside from the difference in mating reaction, the two types of large-celled clones differed from each other with respect to sporulation and growth habit. Sporulation could be induced (on gypsum slants) in the non-mater clones but not in clones which were capable of mating. Also, the cells of the latter tended to grow in clusters, like those of the small-celled clones. (It should be noted that the clustering tendency is obscured in clones that are flaky.) The non-mater clones consisted for the most part of individual cells or cells with single buds.

An indication of the extent of heterogeneity in parental clones was obtained by the following method. The growth on an entire slant, or a substantial portion of it, was suspended in sterile water and dilutions were plated on nutrient agar. The secondary clones obtained in this way were screened for cell size and mating reaction (table 3). The parental clones which were chosen for examination included four from which isolations had already been made with a micromanipulator and three others, 13, 15, and 30, in which large cells had made their appearance after a period of small-celled growth. For comparison, secondary clones were obtained also from four young clones, 242A, B, C, and D, which were derived from four haploid ascospores of a single ascus and which were almost exclusively small-celled at the time the dilutions were made. It will be seen that the younger clones gave only small-celled secondary clones, all of which possessed the expected mating type. The older clones gave varying proportions of large-celled secondary clones, of which the large majority were of the non-mater type. The preponderance of the latter is the reason for the diminished capacity for zygote production which was referred to earlier as characteristic of clones which have become heterogeneous for cell size.

Genetics Tests of Large-Celled Clones: The Non-maters.—As mentioned earlier, the cells of clones of this type can be induced to sporulate. Asci were obtained from ten secondary clones of clone 21, two of clone 27, and one of clone 260. Twenty-three asci were dissected and twenty-one of these exhibited a 2:2 segregation for a and α. One of the secondary clones of clone 21 (21-4), from which six asci were dissected, gave two exceptional asci in which all four spores produced clones that were non-maters. This segregation is characteristic of the tetraploid condition and it is possible that diploidization occurred in 21-4 to produce some cells of composition $aa\alpha\alpha$. An alternative explanation is given below. In any case, it is evident that mutation at the mating-type locus has occurred in these strains and has resulted in the production of $a\alpha$ cells, incapable of mating. The occurrence

of mutation affecting mating type has been observed also by Ahmad[7] in *Saccharomyces cerevisiae* and by Leupold[8] in *Schizosaccharomyces Pombe*.

The occurrence of $a\alpha$ cells was demonstrated in four other parental clones. Clone 1, which consisted largely of cells that were non-maters (table 3), and clones 23, 24, and 25, from which isolations had not been made but which were known to contain large cells, were induced to sporulate on gypsum. Twenty-six four-spored asci were dissected and in all but one a 2:2 segregation for mating type was obtained. In the exceptional ascus, from clone 1, one of the four spores yielded a large-celled clone which was a non-mater. The sporulation of this clone resulted in asci with a 2:2 segrega-

TABLE 4

RESULTS OF CROSS BETWEEN TWO LARGE-CELLED CLONES

4L-II (*aαggFF*) and 260L-V (*aaGGff*)

The clones obtained from the spores of ascus 307 were tested further and the genotype of each spore is given in column 5. Rapid galactose fermenters and flaky clones are indicated by +, slow fermenters and non-flaky clones by −; n indicates the non-mater.

ASCUS	MATING TYPE	GALACTOSE	FLAKY	CONFIRMED GENOTYPE
307A	n	+	+	*aaGGFf*
B	n	−	+	*aaggFf*
C	n	−	+	*aaggFf*
D	n	+	+	*aaGGFf*
308A	α	+	+	
B	a	+	+	
C	n	+	+	
D	n	+	+	
309A	n	+	+	
B	α	+	+	
C	a	+	+	
D	n	−	+	
336A	n	+	−	
B	n	+	+	
C	n	+	+	
D	n	+	−	

tion for a and α. Thus the ascospore from which the clone was grown was probably diploid; the three sister spores, however, gave haploid clones of which two were a and one was α.

We have encountered exceptional asci of this type in other experiments and have adopted as a working hypothesis the explanation that the anomalous segregations are due to postmeiotic nuclear fusion prior to the formation of spores. It is necessary also to assume an extra division of one or more of the nuclei following meiosis. According to this hypothesis, the exceptional ascus is accounted for by assuming that one of the a nuclei divided again after meiosis and that nuclear fusion occurred between an α nu-

cleus and the extra *a* nucleus. The two exceptional asci from 21-4 already described can be accounted for similarly, by assuming an extra division for all four nuclei, with subsequent fusions resulting in four $a\alpha$ spores.

Genetic Tests of Large-Celled Clones: The Maters.—The genetic constitution of the large-celled secondary clones which exhibited a mating reaction was determined from the results of crosses with each other and with haploid clones of known genotype. Eight clones of mating type **α**, isolated from parental clones 1, 4, and 13, and four of mating-type **a**, isolated from 21 and 260, were tested in this manner. A full account of the results of these crosses will be published elsewhere; for the purposes of the present report, the results may be summarized as follows: crosses between large-celled clones gave tetraploid segregations; crosses between large-celled and small-celled clones gave triploid segregations; and crosses between small-celled clones gave diploid segregations. The result of one of the crosses between large-celled clones is given in table 4 as an example. Without exception, the large-celled clones tested in the crosses above have turned out to be diploid and homozygous for one or the other of the mating-type alleles.

Variability of Cell Size in Diploid Clones.—It was mentioned earlier that the cells of diploid clones are not uniformly large; rather, the cells range in size and some are no larger than the cells in haploid clones. As suggested above, the smaller cells may be young cells which have not yet attained full size or mature cells which have their growth inhibited for some reason. The possibility suggested itself, on the other hand, that the vegetative diploid yeast cell is dicaryotic and that the smaller cells are haploid, obtained as a result of the distribution of only one of the two nuclei to the bud. Against this interpretation is the evidence that zygotes are not seen in $a\alpha$ clones nor do such clones mate with either **a** or **α** strains; therefore, free *a* and α cells are not formed in $a\alpha$ clones. Moreover, diploid clones homozygous for mating type have yielded tetraploid segregations only when intercrossed, and thus do not appear to contain haploid cells.

A further test was made with three diploid clones chosen for their relatively high frequency (roughly 50%) of smaller cells. The clones were of composition $a\alpha$, aa, and $\alpha\alpha$, respectively, and were derived from three spores of a four-spored ascus from a cross between two diploids. Samples of the clones were plated out and 45 secondary clones were obtained from the $a\alpha$ clone, 63 from the aa clone, and 60 from the $\alpha\alpha$ clone. All of the secondary clones were of the large-celled type and each gave the mating reaction of the clone from which it was derived. It may be concluded, therefore, that diploid clones do not contain haploid cells in the material under investigation.

Discussion.—The foregoing results indicate the need for caution in interpreting the results of crosses, particularly when mass matings are made, as by the Lindegren technique.[9] In this technique, portions of two clones of opposite mating type are mixed in nutrient broth. After a sufficient time

has elapsed for the production of zygotic cells, the mixture is placed on sporulation medium; the asci thus obtained are dissected and clones are derived from the ascospores for phenotypic test. The occurrence of diploidization in the parental clones, themselves originally haploid, could lead to the production of triploid ($aa\alpha$ or $a\alpha\alpha$ or both) and tetraploid zygotes in the mixture. Also, the parental clones may already contain $a\alpha$ cells at the time the cross is made, as a result of mutation at the mating-type locus. Thus, in addition to the diploid type expected from the cross, five other ascogenous types may be present in the mixture, each of which could yield segregations that would be quite unexpected if the occurrence of diploidization were not recognized.

The effects of diploidization can be largely avoided if matings are made between clones in which large cells, indicating the diploid condition, have not begun to accumulate in sizable numbers. In our material, clones which have been started from haploid ascospores or vegetative cells and incubated for 3–4 days on nutrient agar are generally quite free of diploid cells. In older clones, particularly after a series of subcultures, diploid cells often accumulate at a rapid rate and in some cases virtually replace the haploid cells. The changeover to the diploid condition is frequently indicated by a reduction in the capacity of the clone to exhibit a mating reaction, owing to the large proportion of $a\alpha$ cells which are commonly present in such clones.

The occurrence of an array of ascogenous types can be avoided also if the cells which sporulate are derived from a single zygote. This is accomplished in the methods of Winge and Laustsen[5] and Chen,[10] in which spores or vegetative cells are mated in pairs to produce individual zygotes from which clones are grown for subsequent sporulation. The ploidy of the zygotes that are thus produced will depend of course on the composition of the cells from which they were derived. Single cells, presumed to be diploid, can also be isolated after a mass mating,[11] but here again care must be taken in the selection of clones for mating, to avoid the possibility of selecting cells which were present in the parental clones prior to the cross. This possibility can also be avoided, if the parental clones carry appropriate nutritional markers, by the prototroph-selection method.[12]

The occurrence of diploid cells homozygous for mating type raises the question of whether the strains utilized in this investigation can be regarded as truly heterothallic. If these diploid cells are due to the fusion of haploid cells of like mating type, the strains would seem to possess a tendency toward homothallism. The difference between heterothallism in *S. cerevisiae* and homothallism in *S. chevalieri* is determined by a single gene, at a locus other than that of the mating-type alleles.[13] In *Schizosaccharomyces Pombe*, the homothallic and heterothallic conditions are interconvertible, as a result of mutation at the mating-type locus.[8] Hence, simple genetic

changes can result in changes from one mating system to the other. However, the diploid homozygotes found in our strains differ from those of homothallic strains in two respects: (1) they are capable of forming zygotes with other cells, and (2) they appear to be incapable of producing asci. From the standpoint of reproduction, therefore, the strains under investigation are heterothallic and require the heterozygous condition at the mating-type locus for spore-production.

The results presented above also reveal the manifold effects of the mating-type locus. The production of zygotes, the capacity to sporulate, and the growth habit (whether in small clusters or as free cells) all appear to be dependent on this locus. It would seem reasonable to assume that these characters are manifestations of an effect of the mating-type alleles on the cell surface. If so, it may be possible to alter these characters and, of especial interest, to affect the mating system and sporulation by an appropriate treatment of the cell surface.

Summary.—In the heterothallic strains of Saccharomyces under investigation, diploidization occurs in clones derived from haploid cells The diploid cells are of two types, either homozygous or heterozygous at the mating-type locus. The heterozygotes differ from the homozygotes in growth habit, mating reaction, and in the capacity to sporulate. The heterozygotes, which are capable of sporulation, are incapable of fusing with other cells. The homozygotes do not sporulate; they will fuse, however, with haploid and diploid cells of opposite mating type to produce triploid and tetraploid zygotes.

* This investigation was supported by grants from the Biological and Medical Research Fund of the State of Washington, and The National Institutes of Health, Public Health Service, Grant No. E-328.

[1] Roman, H., Hawthorne, D. C., and Douglas, H. C., PROC. NATL. ACAD. SCI., **37**, 79–84 (1951).

[2] Lindegren, C. C., and Lindegren, G., *J. Gen. Microbiol.*, **5**, 885–893 (1951).

[3] The galactose fermentation tests were conducted with Durham tubes in 10 cc. of a medium containing 2% peptone, 1% yeast extract, and 2% galactose. A clone was classified as a "rapid" fermenter if, starting with an inoculum at the end of a needle, a tubeful of gas was produced within 48 hours at 30°C. A "slow" fermenter produced at most a small bubble of gas during this period.

[4] Tests for the flaky character were conducted in a medium containing DIFCO Yeast Nitrogen Base plus 1% glucose. A small inoculum was added to 10 cc. of this medium and the culture was incubated for a day at 30°C. The standing tubes were then shaken vigorously. In flaky cultures, the cells adhered in visible flakes that quickly settled to the bottom of the tube; in non-flaky cultures, the cells were distributed uniformly in the medium and settled slowly.

[5] Winge, O., and Laustsen, O., *C. R. Lab. Carlsberg, Ser. Physiol.*, **22**, 99–116 (1937).

[6] The nutrient agar used for droplets, slants, and plates, contained 2% peptone, 1% yeast extract, 2% glucose, and 1.5% agar. All incubations were at 30°C.

[7] Ahmad, M., *Nature*, **170**, 546–547 (1952).

[8] Leupold, U., *C. R. Lab. Carlsberg, Ser. Physiol.*, **24**, 381–480 (1950).

[9] Lindegren, C. C., and Lindegren, G., PROC. NATL. ACAD. SCI., **29**, 306–308 (1943).
[10] Chen, S. Y., *C. R. Acad. Sci.*, **230**, 1897–1899 (1950).
[11] Fowell, R. R., *J. Inst. Brewing*, **57**, 180–195 (1951).
[12] Pomper, S., and Burkholder, P. R., PROC. NATL. ACAD. SCI., **35**, 456–464 (1949).
[13] Winge, O., and Roberts, C., *C. R. Lab. Carlsberg, Ser. Physiol.*, **24**, 341–346 (1949).

ASSUMPTION OF BOUNDARY VALUES AND THE GREEN'S FUNCTION IN THE DIRICHLET PROBLEM FOR THE GENERAL LINEAR ELLIPTIC EQUATION

BY FELIX E. BROWDER

DEPARTMENT OF MATHEMATICS, BOSTON UNIVERSITY

Communicated by M. H. Stone, January 6, 1953

During the past year, the author has presented in two notes in these PROCEEDINGS a general method of solving boundary value problems for linear elliptic differential equations of arbitrary order which combines essential features of the integral equation method and Dirichlet's Principle or the method of orthogonal projection.[1] In particular we have shown that the Fredholm alternative holds for the solutions of the Dirichlet problem in an arbitrary bounded domain where the boundary values are to be taken on in the generalized sense familiar from the work of Courant, Friedrichs, and others on the variational approach to elliptic differential equations.[2] It is the first object of this note to outline the proof in the case of two independent variables that for a smoothly bounded domain, smooth boundary data, and an equation with strongly differentiable coefficients, the solution of the Dirichlet problem actually assumes the full set of boundary data in the classical sense. Similar results in n independent variables are derived by the same method. However, for two independent variables, no *a priori* smoothness conditions on the solution need be imposed which are not derived in the general existence theorems.[3] Our second object is to establish the existence and properties of the Green's function for an arbitrary bounded domain in E^n on which the Dirichlet problem has no null solutions.

The proof of Theorem 1 relies extensively upon the results and methods developed by Fritz John in his construction of a fundamental solution in the small.[4] Our methods are extensible without any alteration in principle to the more general case of strongly elliptic systems of differential equations.[5]

Throughout the following discussion, we shall use the notation of our previous notes.

1. Let D be a bounded domain[6] in E^n, $C_c^{2m}(D)$ the linear space of $2m$-

times continuously differentiable functions with compact support in D. $\|f\|_m^2$ is the sum of the L^2-norms of the mth derivatives of f. $C_c^{2m}(D)$ may be completed with respect to $\|\ \ \|_m$ to a (real) Hilbert space $H_m(D)$ and the identity mapping of $C_c^{2m}(D)$ into $L^2(D)$ may be extended to a one-one continuous mapping of $H_m(D)$ onto a linear subset of $L^2(D)$. We shall identify $H_m(D)$ with this subset. A function $u \in L^2(D)$ is said to vanish on the boundary of D together with all its derivatives of order less than m in the variational sense if $u \in H_m(D)$.

Let K be a linear elliptic differential operator on D, $\eta \in C^1(D) \cap L^2(D)$, $g \in C^m(D)$ with square summable mth derivatives in D. The Dirichlet problem for K on D with inhomogeneous part η and boundary function g asks for a function $u \in C^{2m}(D)$ which is a solution of the equation $Ku = \eta$ in D while $u - g$ and all its derivatives of order less than m vanish on the boundary of D in the variational sense.

K is said to have strongly differentiable coefficients on D if the coefficients of the derivatives of order j in K belong to $C^{2m+j+4}(D_1)$ for some domain D_1 containing the closure of D.

$$(N_\epsilon(x_0) = \{x \mid d(x, x_0) < \epsilon\};\ S_\epsilon(x_0) = \{x \mid d(x, x_0) = \epsilon\}).$$

Let $\bar{K}$ be the differential operator on D formally adjoint to K. Then $\bar{K}$ has strongly differentiable coefficients on D if and only if K does.

THEOREM 1.[7] *Let K be a linear elliptic differential operator of order $2m$ having strongly differentiable coefficients on the domain D of E^n. Suppose that u is a measurable function, summable on every compact subset of D, $\eta \in C^1(D)$, and that for every $\varphi \in C_c^{2m}(D)$, $\int_D u\bar{K}\varphi\, dx = \int_D \eta\varphi dx$. Then:*

(1) $u = v$ a. e. in D where $v \in C^{2m}(D)$, $Kv = \eta$. In particular if $u \in C(D)$, $u = v$.

(2) Suppose $u \in C(D)$, $x_0 \in D$, $0 < \epsilon < d(x_0, E^n - D)$. There exists $\epsilon_0 > 0$, independent of x_0, such that for $\epsilon < \epsilon_0$ we may find functions $a_{i_1 \cdots i_r}(x, z)$, $b_{i_1 \cdots i_r}(x, z)$ (for each set of indices in which $r \leq m - 1$) defined for $x \in N_{\epsilon/2}(x_0)$, $z \in N_\epsilon(x_0)$ such that:

(a)
$$|a_{i_1 \cdots i_r}(x, z)| \leq M\epsilon^{-n-r}$$
$$|b_{i_1 \cdots i_r}(x, z)| \leq M\{[d(x, z)]^{2m-n-r} + \epsilon^{2m-n-r}\}$$

where M is independent of x, z, ϵ, and x_0.

(b) For $x \in N_{\epsilon/2}(x_0)$, $r \leq m - 1$,

$$\frac{\partial^r \mu(x)}{\partial x_{i_1} \cdots \partial x_{i_r}} = \int_{N_\epsilon(x_0)} a_{i_1 \cdots i_r}(x, z)u(z)\, dz + \int_{N_\epsilon(x_0)} b_{i_1 \cdots i_r}(z, z)\eta(x)\, dz.$$

Proof: By an extension of the results of F. John,[4] it may be shown that if the linear elliptic differential operator K has strongly differentiable coefficients on D, there exists $\epsilon_0 > 0$ such that for every $x_0 \in \bar{D}$, we may

construct a fundamental solution $e(x, z)$ of $\bar{K}$ for $N_{\epsilon_0}(x_0)$ satisfying the following conditions: (α), $\varphi(x) = \int_{N_{\epsilon_0}(x_0)} e(x, z) K\varphi(z)\, dz$ if $\varphi \in C_c^{2m}(N_{\epsilon_0}(x_0))$; (β), $\dfrac{\partial^{r+s} e(x, z)}{\partial x_{i_1} \ldots \partial x_{i_r}\, \partial z_{j_1} \ldots \partial z_{j_s}}$ exists for all sets of indices for which $r \leq m - 1$, $s \leq 2m$ for $x \neq z$, is continuous in $N_{\epsilon_0}(x_0)$ for $x \neq z$ and satisfies the inequality

$$\left| \frac{\partial^{r+s} e(x, z)}{\partial x_{i_1} \ldots \partial x_{i_r}\, \partial z_{j_1} \ldots \partial z_{j_s}} \right| \leq M\{[d(x, z)]^{2m-n-r-s} + |\log d(x, z)|\}$$

where M is independent of x, z, and x_0; (γ), there exist functions $a(x)$, $c(x, z)$ defined for $x, z \in N_{\epsilon_0}(x_0)$, $a(x)$ bounded on $N_{\epsilon_0}(x_0)$, $c(x, z)$ and its derivatives of order less than $2m - 1$ with respect to the z-variables existing and summable in x on $N_{\epsilon_0}(x_0)$, for which

$$\frac{\partial(e(x,z))}{\partial z_i} = -\frac{\partial(e(x, z))}{\partial x_i} + a(x)e(x, z) + c(x, z).$$

Let $\eta \in C^1(\bar{N}_{\epsilon_0}(x_0))$, $\psi(z) = \int_{N_{\epsilon_0}(x_0)} \eta(x)e(x, z)\, dx$. Then $\psi \in C^{2m}(N_{\epsilon_0}(x_0))$, $\bar{K}\psi = \eta$ on $N_{\epsilon_0}(x_0)$. For by (β), $\psi \in C^{2m-1}(N_{\epsilon_0}(x_0))$ and by (γ),

$$\frac{\partial \psi}{\partial z_i}(z) = \int_{N_{\epsilon_0}(x_0)} \eta(x) \frac{\partial e(x, z)}{\partial z_i}\, dx = -\int_{N_{\epsilon_0}(x_0)} \eta(x) \frac{\partial e(x, z)}{\partial x_i}\, dx + \int_{N_{\epsilon_0}(x_0)} a(x)\eta(x)e(x, z)\, dx + \int_{N_{\epsilon_0}(x_0)} c(x, z)\, \eta(x)\, dx.$$

Since $\eta \in C^1(\bar{N}_{\epsilon_0}(x_0))$,

$$+ \int_{N_{\epsilon_0}(x_0)} \eta(x) \frac{\partial(e(x, z))}{\partial x_i}\, dx = -\int_{N_{\epsilon_0}(x_0)} \frac{\partial \eta}{\partial x_i} e(x, z)dx - \int_{S_{\epsilon_0}(x_0)} \eta(x)e(x, z) \cos(n, x_i)\, dSx.$$

Thus $\dfrac{\partial \psi}{\partial z_i} \in C^{2m-1}(D)$ and $\psi \in C^{2m}(D)$. If $\varphi \in C_c^{2m}(D)$, by (α),

$$\int_{N_{\epsilon_0}(x_0)} \varphi(z)\bar{K}\psi(z)\, dz = \int_{N_{\epsilon_0}(x_0)} K\varphi(z) \left\{ \int_{N_{\epsilon_0}(x_0)} \eta(x)e(x, z)\, dx \right\} dz = \int_{N_{\epsilon_0}(x_0)} \eta(x) \left\{ \int_{N_{\epsilon_0}(x_0)} K\varphi(z)e(x, z)\, dz \right\} dx = \int_{N_{\epsilon_0}(x_0)} \eta(x)\varphi(x)\, dx$$

Therefore $\bar{K}\psi = \eta$.

Given any $\epsilon > 0$, there exists $\rho \in C_c^{4m}(N_{\epsilon/2}(0))$, $M_1 > 0$, such that $\left| \dfrac{\partial^r \rho}{\partial x_{i_1} \ldots \partial x_{i_r}} \right| \leq M_1 \epsilon^{-r}$ for $0 \leq r \leq 4m$. If $x_0 \in D$, $0 < \epsilon < \min(\epsilon_0, d(x_0, E^n - D))$, let $e_1(x, z) = \rho(x - z)e(x, z)$, $e_2(x, z) = e(x, z) - e_1(x, z)$. If $\varphi \in C_c^{2m}(N_{\epsilon/2}(x_0))$, $\psi(z) = \int_{N_{\epsilon/2}(x_0)} e_1(x, z)\varphi(x)\, dx$, then $\psi \in C_c^{2m}(N_\epsilon(x_0))$. By our hypothesis that u is a weak solution of $Ku = \eta$,

$$\int_{N_\epsilon(x_0)} \eta(z)\psi(z)\,dz = \int_{N_\epsilon(x_0)} u(z)\bar{K}\psi(z)dz =$$
$$\int_{N_\epsilon(x_0)} u(z)\{\varphi(z) - \int_{N_\epsilon(x_0)} \bar{K}_z e_2(x, z)\varphi(x)\,dx\}dz.$$

If we set $e_3(x, z) = \bar{K}_z e_2(x, z)$,

$$\int_{N_\epsilon(x_0)} \eta(z)\psi(z)\,dz = \int_{N_\epsilon(x_0)} \varphi(x)\{\int_{N_\epsilon(x_0)} e_1(x, z)\eta(z)\,dz\}dx =$$
$$\int_{N_\epsilon(x_0)} \varphi(x)\{u(x) - \int_{N_\epsilon(x_0)} e_3(x, z)u(z)\,dz\}\,dx.$$

If $v_{x_0}(x) = \int_{N_\epsilon(x_0)} e_3(x, z)u(z)\,dz + \int_{N_\epsilon(x_0)} e_1(x, z)\eta(z)\,dz$ then $u = v_{x_0}$ a. e. in $N_{\epsilon/2}(x_0)$ and (1) follows easily. If u is continuous, $u(x) = \int_{N_\epsilon(x_0)} e_3(x, z)u(z)\,dz + \int_{N_\epsilon(x_0)} e(x, z)\eta(z)\,dz - \int_{N_\epsilon(x_0)} e_2(x, z)\eta(z)\,dz$. If we differentiate with respect to the x-variables and apply (β) and the inequalities on ρ and its derivatives, (2) follows. Q. E. D.

THEOREM 2. *Suppose K is a linear elliptic differential operator with strongly differentiable coefficients on the bounded domain D in E^2. Suppose that the boundary of D is locally connected and has no isolated points. Let $g \in C^{2m+1}(\bar{D})$, $\eta \in C^1(D)$ and uniformly bounded on D. Then if u is the solution of the Dirichlet problem for K on D with boundary function g and inhomogeneous part η, $u - g$ and all its derivatives of order less than m approach zero as their arguments approach a boundary point. In particular u and all its derivatives of order less than m are continuous on $\bar{D}$.*

Proof: If $w = u - g$, then $w \in H_m(D)$ and $Kw = \eta - Kg$. It follows from Theorem 1 that

$$\left|\frac{\partial^r w(x_0)}{\partial k_{i_1} \ldots \partial k_{i_r}}\right| \leq \epsilon^{-n-r} \int_{N_\epsilon(x_0)} |w(x)|\,dx + M\{\epsilon^{2m-r} + \epsilon\}$$

for $r \leq m - 1$.

LEMMA 1.[3] *Suppose that the domain D in E^2 has a locally connected boundary without isolated points. If $w \in H_m(D)$, $x_0 \in D$, $h = d(x_0, E^n - D)$, then $\int_{N_{h/2}(x_0)} |w(x)|\,dx \leq Ch^{m+1}k(h, w)$, where C is independent of h and $k(h, w) \to 0$ as $h \to 0$.*

Our conclusion follows from Lemma 1. Q. E. D.

THEOREM 3. *Let K be a linear elliptic differential operator on the bounded domain D of E^n. Suppose that the boundary of D has a continuously turning normal vector, $g \in C^{2m+1}(\bar{D})$, $\eta \in C^1(D)$ and uniformly bounded on D. Let u be the solution of the Dirichlet problem for K on D with boundary function g and inhomogeneous part η. If $u \in C^{m-1}(\bar{D})$, then $u - g$ and all its derivatives of order less than m approach zero as their argument approaches the boundary of D.*

2. Let K be a suitably differentiable linear elliptic differential operator of order $2m$ on the bounded domain D of E^n. $G(x, z)$, defined and measurable on $D \times D$ and summable in x for each fixed z in D, is said to be a

Green's function for K on D if for $\eta \in C^1(D) \cap L^2(D)$, $\psi(z) = \int_D G(x, z)\eta(x)\,dx$, we have $\psi \in C^{2m}(D) \cap H_m(D)$, $K\psi = \eta$ in D.

By our previous results, a necessary condition that there exist a Green's function for K on D is that the Dirichlet problem for K on D have no null solutions, and if a Green's function exists, it is unique. The converse is shown in the following:

THEOREM 4. *Let K be a linear elliptic differential operator of order $2m$ on the bounded domain D of E^n with suitably differentiable coefficients on a domain containing $\bar{D}$. Suppose D is a domain whose boundary has n-dimensional measure zero. Suppose further that the Dirichlet problem for K on D has no non-trivial null solutions. Then there exists a Green's function for K on D which is $2m$-times continuously differentiable in x and z for $x \neq z$ and is a fundamental solution for K on D.*

Proof: Under the assumption on the differentiability of the coefficients of K and on the measure of the boundary of D, there exists a domain D_1 on which K is suitably differentiable, $\bar{D} \subset D_1$, and for which the Dirichlet problem for K has no non-trivial null solutions. If D_2 is another domain in E^n, $\bar{D} \subset D_2 \subset \bar{D}_2 \subset D_1$, the family $\{K\varphi \mid \varphi \in C^{2m}(D_2)\}$ is dense in $L^2(D_2)$. By a variant of the method of Levi, there exists a fundamental solution $e(x, z)$ for K on D_2 which, by Theorem 1, for fixed x in D_2, is $2m$-times continuously differentiable in z for $z \neq x$. For $x \in D$, let $e_0(x, z) = e(x, z)\sigma(x, z)$ where σ is $2m$-times continuously differentiable in z and vanishes for z on some neighborhood of x contained in D. The Dirichlet problem for $Ku = 0$ on D with boundary function $e_0(x, z)$ has a unique solution $h(x, z) \in C^{2m}(D)$ for each fixed x in D. Let $G(x, z) = e(x, z) - h(x, z)$. If $\sigma(x, z)$ is chosen so that it, its derivatives, and the radius of the neighborhood of x in which it vanishes are suitably continuous in x, then $G(x, z) = p(x, z) + s(x, z)$, $p(x, z)$ continuous in x and vanishing for each x on a neighborhood of the boundary of D, $h(x, z) \in H_m(D)$ for each fixed x in D and varying continuously in $H_m(D)$ with x in D. $G(x, z)$ is a fundamental solution for K on D and is $2m$-times continuously differentiable in z for each x in D.

Let $\eta \in C^1(D) \cap L^2(D)$, $\psi(z) = \int_D G(x, z)\eta(x)\,dx$. Suppose $\varphi \in C_c^{2m}(D)$. Then

$$\int_D \bar{K}_z\varphi(z)\psi(z)\,dz = \int_D \bar{K}_z\varphi(z)\left\{\int_D G(x, z)\eta(x)\,dx\right\}dz = \int_D \eta(x)\left\{\int_D G(x, z)\bar{K}_z\varphi(z)\,dz\right\}dx = \int_D \eta(x)\varphi(x)\,dx.$$

Since $\eta \in C^1(D)$, by Theorem 1, $\psi \in C^{2m}(D)$, $K\psi = \eta$. If $\eta \in C_c^1(D)$, it is easily seen that $\psi \in H_m(D)$. It follows from our previous results that there exists a continuous linear transformation T from $L^2(D)$ into $H_m(D)$ such that

$$\int_D T\eta(x)\bar{K}\varphi(x)\,dx = \int_D \eta(x)\varphi(x)\,dx,$$

$\eta \in L^2(D)$, $\varphi \in C_c^{2m}(D)$. For $\eta \in C_c^1(D)$, $T\eta = \psi$, by Theorem 1. If $\eta_i \to \eta$ in $L^2(D)$, η_i converging uniformly to η on every compact subset of D, $\eta_i \in C^1(D)$, $\eta \in C^1(D) \cap L^2(D)$, then $T\eta_i \to T\eta$ in $H_m(D)$ and consequently in $L^2(D)$. But $T\eta_i(z) = \int_D G(x, z)\eta_i(x)\,dx \to \int G(x, z)\eta(x)\,dx$ for every z in D. Therefore $T\eta = \psi$ a. e. in D. But $\psi \in C^{2m}(D)$, and $T\eta \in C^{2m}(D) \cap H_m(D)$. Therefore $\psi = T\eta$ in D and $\psi \in H_m(D)$.

Let $\bar{G}(x, z)$ be the Green's function for $\bar{K}$ on D, which may be constructed by the same process since the Dirichlet problem for $\bar{K}$ can have no non-trivial null solutions on D under our hypotheses. $\bar{G}(x, z)$, like $G(x, z)$, is $2m$-times continuously differentiable in z for fixed x in D.

If $w, y \in D$, $\epsilon < \min\{d(w, E^n - D), d(y, E^n - D), d(w, y)/2\}$, let $D_\epsilon = D - N_\epsilon(w) - N_\epsilon(y)$. Then (if M is a bilinear concomitant of K, $\bar{M}$ of $\bar{K}$, with respect to the exterior of D),

$$\begin{aligned} 0 &= \int_D \bar{G}(w, z)K_z G(y, z)\,dz - \int_D G(y, z)\bar{K}_z G(w, z)\,dz \\ &= -\int_{S_\epsilon(w)} M_z(G(w, z), \bar{G}(y, z))dS_z - \int_{S_\epsilon(y)} M_z(G(w, z), \bar{G}(y, z))\,dS_z \\ &= -\int_{S_\epsilon(w)} M_z(G(w, z), \bar{G}(y, z))\,dS_z + \int_{S_\epsilon(y)} \bar{M}_z(\bar{G}(w, z), G(y, z))\,dS_z \\ &= -\bar{G}(y, w) + G(w, y). \end{aligned}$$

Thus $G(x, z) = \bar{G}(z, x)$ is $2m$-times continuously differentiable in x for fixed z in D. (Analyticity in both variables may be established similarly for equations with analytic coefficients.) Q. E. D.

[1] Browder, F. E., (I) "The Dirichlet Problem for Linear Elliptic Equations of Arbitrary Even Order with Variable Coefficients," PROC. NATL. ACAD. SCI., **38**, 230–235 (1952); (II) "The Dirichlet and Vibration Problems for Linear Elliptic Differential Equations of Arbitrary Order," *Ibid.*, **38**, 741–747 (1952).

[2] Cf. Courant-Hilbert, *Methoden der Mathematische Physik*, v. 2, Berlin, 1937, Chap. VII.

[3] Cf. reference 2, pp. 495–497.

[4] John, F., "General Properties of Solutions of Linear Elliptic Partial Differential Equations," Proceedings of the Symposium on Spectral Theory and Differential Problems, Stillwater, Okla. (1951).

[5] A detailed exposition of our results for this case will appear in the near future. The definition of strictly elliptic systems was given by M. I. Visik, *Doklad. Akad. Nauk. S.S.S.R. (N.S.)*, **74**, 881–884 (1950).

[6] The results and proofs are local in character. Boundedness is assumed for the sake of simplicity of statement.

[7] Theorem 2 of (I) and Theorem 5 of (II) are immediate corollaries of Theorem I.

LINEAR PARABOLIC DIFFERENTIAL EQUATIONS OF ARBITRARY ORDER; GENERAL BOUNDARY-VALUE PROBLEMS FOR ELLIPTIC EQUATIONS

BY FELIX E. BROWDER

DEPARTMENT OF MATHEMATICS, BOSTON UNIVERSITY

Communicated by M. H. Stone, January 6, 1953

The close relation between the theories of elliptic and parabolic partial differential equations has long been familiar to mathematicians who have worked with either. The simplest instances appear in the method of separation of variables and in the fact that limits as $t \to +\infty$ of solutions of parabolic equations, if such limits exist, must be solutions of elliptic equations. The first purpose of this note is to substantiate this relationship by deriving from the author's results on linear elliptic equations,[1] a general existence and differentiability theorem for the solutions of the mixed initial-boundary value problem for a broad class of linear parabolic equations of arbitrary order.[2] Our second purpose is to outline some theorems concerning a more general class of boundary value problems for linear elliptic equations of arbitrary order established by the author previously for the Dirichlet problem—the Fredholm alternative, complete continuity of the inverse of the differential operator, completeness of the eigenfunctions for self-adjoint operators, semi-boundedness, and existence of a "Green's function" integral operator representation for the inverse. Our results may be generalized to a large class of elliptic systems of differential equations. A detailed exposition of our results for this general case will appear in the near future.[3]

1. Let D be a bounded domain in E^n, K a suitably differentiable linear elliptic differential operator on D. Let $R = \{t \mid t > 0\}$.

If $u(x, t)$ defined on $D \times R$ is $2m$-times differentiable with respect to the x-variables and once differentiable with respect to t, u satisfies the linear parabolic equation of order $2m$ corresponding to K with inhomogeneous part η if

$$(-1)^{m+1}Ku + \eta = \frac{\partial u}{\partial t} \qquad \text{for } (x, t) \in D \times R, \eta \in C^1(D) \cap L^2(D).$$

Such a function u is said to satisfy the initial-boundary value problem with initial function h and boundary function g if $u(x, t) - g$ belongs to $H_m(D)$ for every $t > 0$ and if $u(x, t) \to h(x)$ in $L^2(D)$ as $t \to 0$. If the Dirichlet problem for K on D has no non-trivial null solutions, then by introducing the new dependent variable $u_1 = u - w$, where w is the solution of the Dirichlet problem for K on D with boundary function g and inhomogeneous part $(-1)^m \eta$, we obtain a parabolic equation with

zero inhomogeneous part and boundary function and with the initial function h replaced by $h - g$. We shall restrict ourselves to this latter case.

THEOREM 1. *Let K be a suitably differentiable, self-adjoint, linear elliptic differential operator on the bounded domain D of E^n whose boundary has zero n-measure. Suppose the Dirichlet problem for K on D has no null solutions. Then there exists a solution $u(x, t)$ of the parabolic equation $(-1)^{m+1}Ku = \partial u/\partial t$ on $D \times R$ for boundary function zero and arbitrary initial function $h \in L^2(D)$ with $u(x, t)$ $2m$-times continuously differentiable in x and infinitely differentiable in t for $(x, t) \in D \times R$. If the coefficients of K are analytic in D, u is analytic in x.*

LEMMA 1. *Let K be a suitable differentiable linear elliptic differential operator on the bounded domain D of E^n whose boundary has zero n-measure. Suppose that the Dirichlet problem for K on D has no non-trivial null solutions. Let $G(x, z)$ be the Green's function of K on D.*[4] *Then the function $G_r(x, z)$ defined recursively for $x, z \in D$, $x \neq z$, by the relation $G_r(x, z) = \int_D G(x, y)\, G_{r-1}(y, z)\, dy$, $G_1 = G$, exists for every positive integer r and satisfies the following conditions:*

$(\alpha)_r$. *$G_r(x, z)$ is a fundamental solution for K^r on D.*

$(\beta)_r$. *$G_r(-, z)$, $\partial G_r/\partial z_j\,(-, z)$ for z in D lie in $L^2(D)$ for $r > n/2m$.*

$(\gamma)_r$. *If $\varphi \in C_c^{2m}(D) \cap H_m(D)$, $K\varphi = \lambda\varphi$, then $\lambda^r \int_D G_r(x, z)\varphi(x)dx = \varphi(z)$, $\lambda^r \int_D \partial G_r/\partial z_j(x, z)\varphi(x)dx = \partial\varphi/\partial z_j(z)$.*

$(\delta)_r$. *Let D_1 be a subdomain of D, $\bar{D}_1 \subset D$, $0 < \epsilon < d(D_1, E^n - D)$. Then corresponding to ϵ, $G_r(x, z)$ may be written in the form $s_r(x, z) + h_r(x, z)$ where $s_r(x, z) = 0$ for $d(x, D_1) \geq \epsilon$, $h_r(-, z)$ for fixed z lies in $H_m(D)$, $\int_D |s_r(x, z) - s_r(x, z_1)|\, dx \to 0$ as $z \to z_1$ in D_1 and similarly for $\partial s_r(x, z)/\partial z_j$. There exists $c_r > 0$, for which*

$$|s_r(x, z)|, \left|\frac{\partial s_r(x, z)}{\partial z_j}\right| \leq c_r\{[d(x, z)]^{2mr-n-1} + |\log d(x, z)| + 1\}.$$

The mapping of D_1 into $H_m(D)$ defined by $z_1 \to h_r(-, z_1)$ is continuous on D_1.

Proof: It suffices to prove $(\alpha)_r$, $(\gamma)_r$, $(\delta)_r$ by induction since $(\delta)_r$ implies $(\beta)_r$. The Green's function exists and $(\alpha)_1$, $(\gamma)_1$, $(\delta)_1$ follow from its properties.[4] Suppose $G_{r-1}(x, z)$ exists and $(\alpha)_{r-1}$, $(\gamma)_{r-1}$, and $(\delta)_{r-1}$ have been proved. Then it follows from $(\delta)_{r-1}$ that $G_r(x, z) = \int_D G(x, y)G_{r-1}(y, z)\, dy$ exists for $x, z \in D$, $x \neq z$. By $(\delta)_{r-1}$, $G_{r-1}(x, z) = s_{r-1}(x, z) + h_{r-1}(x, z)$ for $z \in D_1$, the decomposition corresponding to $\epsilon/2$. On the other hand, $G(x, y) = S(x, y) + h(x, y)$, where $h(x, y) \in H_m(D)$ for fixed y in $N_\epsilon(D_1)$, $\int_D |S(x, y)|\, dx$ uniformly bounded for $y \in N_\epsilon(D_1)$, and $S(x, y) = 0$ for $d(x, y) \geq \epsilon$. For $z \in D_1$, $G_r(x, z) = \int_D G(x, y)\, s_{r-1}(y, z)\, dy + \int_D G(x, y)\, h_{r-1}(y, z)\, dy = \int_D G(x, y)\, h_{r-1}(y, z)\, dy + \int_D S(x, y)\, s_{r-1}(y, z)\, dy + \int_D h(x, y)\, s_{r-1}(y, z)\, dy$. If we set $s_r(x, z) = \int_D S(x, y)\, s_{r-1}(y, z)\, dy$, $h_r(x, z) = \int_D G(x, y)\, h_{r-1}(y, z)\, dy + \int_D h(x, y)\, s_{r-1}(y, z)\, dy$, then s_r and h_r will give the decomposition of G_r for D_1 demanded by $(\delta)_r$.

If $\varphi \in C_c^{2m}(D)$, $\int_D G_r(x, z)\bar{K}^r\varphi(z)\,dz = \int_D G(x, y)\{\int_D G_{r-1}(y, z)\bar{K}^{r-1}\bar{K}\varphi(z)\,dz\}\,dy = \int_D G(x, y)\bar{K}\varphi(y)\,dy = \varphi(x)$ and $(\alpha)_r$ follows.

If $\varphi \in C^{2m}(D) \cap H_m(D)$, $K\varphi = \lambda\varphi$, then by the definition of Green's function, $\int_D G(x, z)\varphi(x)\,dx = \varphi(z)/\lambda$. Thus $\int_D G_r(x, z)\varphi(x)\,dx = \int_D\{\int_D G(x, y)\,G_{r-1}(y, z)\,dy\}\varphi(x)\,dx = \int_D G_{r-1}(y, z)\{\int_D G(x, y)\varphi(x)\,dx\}dy = \frac{1}{\lambda}\int_D G_{r-1}(y, z)\varphi(y)\,dy = \varphi(z)/\lambda^r$. The remainder of $(\gamma)_r$ follows. Q. E. D.

Proof of Theorem 1: We have shown previously that the eigenfunctions of the self-adjoint linear elliptic operator K for the Dirichlet problem are complete in $H_m(D)$ and *a fortiori* in $L^2(D)$.[5] After an orthonormal basis has been chosen for the eigenspace corresponding to each eigenvalue, considered as a subset of $L^2(D)$, the resulting set of eigenfunctions $\{\varphi_i\}$ is a complete orthonormal set in $L^2(D)$ and orthogonal in $H_m(D)$ with respect to the inner product on $H_m(D)$ induced from $C_c^{2m}(D)$ by $(f, g) = (-1)^m\int_D Kf\cdot g\,dx + k_0\int_D f\cdot g\,dx$, where the real number k_0 has been chosen so that $K + (-1)^m k_0$ is positive.[5] If $\|f\| = (f, f)^{1/2}$, $\|\varphi_i\| = (-1)^m\lambda_i + K_0 \to +\infty$ as $i \to +\infty$.

By Lemma 1, if $r > n/2m$, $G(-, z)$ for z in D lies in $L^2(D)$ while $\int_D G_r(x, z)\varphi_i(x)\,dx = \varphi_i(z)/\lambda_i^r$. By the completeness of $\{\varphi_i\}$ in $L^2(D)$ and Parseval's equality, $\sum_{i=1}^{\infty}[\varphi_i(z)]^2/\lambda_i^{2r} = \int_D|G_r(x, z)|^2\,dx < \infty$ and $\sum_{i=1}^{\infty}\left[\frac{\partial\varphi_i(z)}{\partial z_j}\right]^2/\lambda_i^{2r} = \int_D\left[\frac{\partial G_r}{\partial z_j}(x, z)\right]^2 dx < \infty$.

If h is the given initial function, $c_i = \int h(x)\varphi_i(x)\,dx$, we set $u_s(x, t) = \sum_{i=1}^{s} c_i\varphi_i(x)\,e^{(-1)^{m+1}\lambda_i t}$. For fixed $t > 0$, $\{u_s(-, t)\}$ is a Cauchy sequence in $L^2(D)$ and in $H_m(D)$. Let $u(-, t)$ be the element of $H_m(D)$ which is the limit of this sequence. $u(-, t) \to h$ in $L^2(D)$ as $t \to 0$. By its definition $u_s(x, t)$ is $2m$-times continuously differentiable in x, infinitely differentiable in t and satisfies the differential equation, $(-1)^{m+1}Ku_s = \partial u_s/\partial t$. If $\theta \in C_c^{2m}(D)$, $(-1)^{m+1}\int_D u_s(x, t)\bar{K}\theta(x)\,dx = \int_D \frac{\partial u_s}{\partial t}(x, t)\theta(x)\,dx$.

For an interior point z of D,

$$\left|\frac{\partial^k u_{s_1}}{\partial t^k}(z, t) - \frac{\partial^k u_s}{\partial t^k}(z, t)\right| = \left|\sum_{i=s+1}^{s_1}(-1)^{k(m+1)}\lambda_i^k\,c_i\varphi_i(z)e^{(-1)^{m+1}\lambda_i t}\right| \leq \left\{\sum_{i=s+1}^{s_1} c_i^2\right\}^{1/2}\left\{\sum_{i=s+1}^{s_1}[\varphi_i(z)]e^{2(-1)^{m+1}\lambda_i t}\lambda_i^{2k}\right\}^{1/2}.$$

for any non-negative integer x. For fixed $t > 0$, since $(-1)^m\lambda_i \to +\infty$, $|\lambda_i^{2k}e^{+2(-1)^{m+1}\lambda_i t}| < 1/\lambda_i^{2r}$ for large i. Thus $\left|\frac{\partial^k u_s}{\partial t^k}(z, t) - \frac{\partial u_{s_1}}{\partial t^k}(z, t)\right| \to 0$

as $s, s_1 \to +\infty$, uniformly in a neighborhood of t bounded away from zero. Similarly $\left| \frac{\partial^{k+1}}{\partial t^k \partial z_j} u_s(z, t) - \frac{\partial^{k+1}}{\partial t \partial z_j} u_{s_1}(z, t) \right| \to 0$ as $s, s_1 \to +\infty$, uniformly for z in the neighborhood of a point z_0 in D and for t in a neighborhood of t_0 bounded away from zero. Therefore $u_s(x, t)$, its derivatives with respect to t, $\frac{\partial}{\partial z_j}\left(\frac{\partial u_s}{\partial t}\right)$, and its derivatives with respect to t converge uniformly to u and its corresponding derivatives, which exist and are continuous in x and t. Therefore $\int_D (-1)^{m+1} u \bar{K}\theta \, dx = \int_D \frac{\partial u}{\partial t} \, dx$ for every $\theta \in C_c^{2m}(D)$ and therefore $u(-, t) \in C^{2m}(D) \cap H_m(D)$ for $t > 0$, $(-1)^{m+1} Ku = (\partial u/\partial t)$. Q. E. D.

2. Let K be a suitably differentiable linear elliptic differential operator of order $2m$ on the bounded domain D of E^n. D is said to be smoothly bounded if its boundary B is locally a $(n-1)$-manifold of class C^{2m} imbedded $2m$-times continuously differentiably in E^n. Following the example of the Dirichlet problem, for the general class of linear boundary-value problems, each problem will be defined with respect to a linear subset $\tilde{H}_b$ of $C^{2m}(\bar{D})$ such that $C_c^{2m}(D) \subset \tilde{H}_b$. If $\{b_\alpha \mid \alpha \in \Omega\}$ is a family of linear mappings from $C^{2m}(D)$ to linear spaces S_α, they will define such a subset $\tilde{H}_b = \{f \mid b_\alpha(f) = 0 \text{ for } \alpha \in \Omega\}$ provided that $b_\alpha(f) = 0$ for $f \in C_c^{2m}(D)$, α in Ω. (If $0 \leq j \leq 2m$, $u \in C^{2m}(\bar{D})$, $\|u\|'_j = \sum_{i \leq} \|u\|_i$.)

K may be written in the form, $K = \sum_{i_1 \ldots i_{2m}} a_{i_1 \ldots i_{2m}}(x) \frac{\partial^{2m}}{\partial x_{i_1} \ldots \partial x_{i_{2m}}} + R$, order of $R < 2m$.

If p is a bilinear form on $C^{2m}(D)$, the boundary-value problem corresponding to $(K, \tilde{H}_b, p)$ is said to be proper if the following conditions hold:

(*A*) For $u, v \in \tilde{H}_b$,

$$(-1)^m \int_D Ku \cdot v \, dx = \sum_{i_1 \ldots i_{2m}} \int_D a_{i_1 \ldots i_{2m}}(x) \frac{\partial^m u}{\partial x_{i_1} \ldots \partial x_{i_m}} \frac{\partial^m v}{\partial x_{i_{m+1}} \ldots \partial x_{i_{2m}}} \, dx + r(u, v) + p(u, v) + s(u, v)$$

such that,

(1) There exists $c > 0$ for which $|r(u, v)| \leq c\|u\|'_{m-1}\|v\|'_m$ when $u, v \in \tilde{H}_b$.

(2) For $u, v \in \tilde{H}_b$, $p(u, v) = p(v, u)$ while $p(u, u) > 0$ if $u \neq 0$.

(3) $s(u, v) = \sum_{\substack{j_1 \ldots j_r \\ k_1 \ldots k_s}} {}_{r, s \leq m-1} \int_B C_{j_1 \cdots j_r k_1 \cdots k_s}(x) \frac{\partial^r u}{\partial x_{j_1} \ldots \partial x_{j_r}} \frac{\partial^s v}{\partial x_{k_1} \ldots \partial x_{k_r}} \, dS.$

(*B*) There exists $d \geq 0$, $\rho > 0$ for which (for $u \in \tilde{H}_b$)

$$\int_D \left\{ \sum_{i_1 \cdots i_{2m}} a_{i_1 \cdots i_{2m}}(x) \frac{\partial^m u}{\partial x_{i_1} \cdots \partial x_{i_m}} \frac{\partial^m u}{\partial x_{i_{m+1}} \cdots \partial x_{i_{2m}}} \right\} dx + p(u, u) \geq \rho \|u\|'^2_m - d\|u\|'^2_{m-1}.$$

Corresponding to a proper boundary-value problem $(K, \tilde{H}_b, p)$ we may define an inner product on $\tilde{H}_b$ by

$$(u, v) = \int_D \left\{ \sum_{i_1 \cdots i_{2m}} a_{i_1 \cdots i_{2m}}(x) \frac{\partial^m u}{\partial x_{i_1} \cdots \partial x_{i_m}} \frac{\partial^m v}{\partial x_{i_{m+1}} \cdots \partial x_{i_{2m}}} \right\} dx + p(u, v) + d[(u, v)'_{m-1}].$$

For $u \in \tilde{H}_b$, $(u, u) > 0$ unless $u = 0$. Setting $\|u\| = (u, u)^{1/2}$, we define H_b as the completion of $\tilde{H}_b$ with respect to $\| \quad \|$. H_b may be identified with a subset of $L^2(D)$ and is described as the space of functions satisfying the boundary-value problem $(K, \tilde{H}_b, p)$ in the variational sense.

$(K, \tilde{H}_b, p)$ is said to be strongly proper if it is proper and if in addition:

(C_1) For each $v \in \tilde{H}_b$ there is a constant c_v such that $\left| \int_D Ku \cdot v \, dx \right| \leq c_v \{ \|u\|'^2_m + p(u, u)^{1/2} \}$ for all $u \in H_b \cap C^{2m}(D)$.

The boundary-value problem (K, H_b, p) is said to be adjoint-proper if it is strongly proper and if:

(C_2) For $u, v \in H_b$, $\int_D Ku \cdot v \, dx = \int_D u \cdot Kv \, dx$.

The Dirichlet problem is the adjoint-proper problem $(K, C_c^{2m}(D), 0)$. Condition (*B*) is satisfied independently of $\tilde{H}_b$ for all linear elliptic equations in two independent variables, all second-order linear elliptic equations in n independent variables, fourth order equations in three independent variables, and more generally whenever the form $\sum_{i_1 \cdots i_{2m}} a_{i_1 \cdots i_{2m}}(x) \xi_{i_1 \cdots i_m} \xi_{i_{m+1} \cdots i_{2m}}$ is uniformly positive definite for x in D. The classical second and third boundary-value problems for second-order equations may be formulated as adjoint-proper boundary-value problems.[6]

THEOREM 2. *Let K be a suitably differentiable linear elliptic differential operator on the smoothly bounded domain D of E^n. Suppose that $(K, \tilde{H}_b, p)$ is a proper boundary-value problem on D. Let $g \in C^{2m+1}(D)$, $\eta \in C^1(D) \cap L^2(D)$. Suppose that there exists no $v \in C^{2m}(D) \cap H_b$ different from zero for which $Kv = 0$. Then for every g, η there exists $u \in C^{2m}(D)$ for which $Ku = \eta$ and $u - g$ lies in H_b. More generally, if $\int \varphi \eta \, dx = 0$ for all solutions φ of the equation $K\varphi = 0$, $\varphi \in H_b$, then there exists $u \in C^{2m}(D) \cap H_b$ such that $Ku = \eta$.*

THEOREM 3. *Suppose that $(K, \tilde{H}_b, p)$ is a strongly proper boundary-value problem. Then the conditions given in Theorem 2 for the existence of solutions are necessary as well as sufficient. The eigenvalues of K with respect to the boundary-value problem are discrete, the corresponding eigenspaces*

finite dimensional. If $\{\lambda_i\}$ is the family of eigenvalues, then $(-1)^m R(\lambda_i) \to +\infty$ as $i \to +\infty$.

THEOREM 4. *If K is self-adjoint and $(K, \tilde{H}_b, p)$ is adjoint-proper, then the eigenfunctions of K with respect to this boundary-value problem are complete in H_b as well as in $L^2(D)$.*

The proofs of these theorems which are analogous to the corresponding proofs of the Dirichlet problem are based upon Theorem 2 of (I) as well as the following lemma:

LEMMA 2. *Let D be a smoothly bound domain, $\{u_i\}$ a family of functions from $C^{2m}(D)$ such that $\|u_i\|'_m \leq M$ for all i. Then there exists an infinite subsequence $\{u_{i_r}\}$ for which $\|u_{i_r} - u_{i_s}\|'_{m-1} \to 0$ as $r, s \to +\infty$.*

[1] Browder, F. E., (I) "The Dirichlet Problem for Linear Elliptic Equations of Arbitrary Order with Variable Coefficients," PROC. NATL. ACAD. SCI., **38**, 230–235 (1952); (II) "The Dirichlet and Vibration Problems for Linear Elliptic Differential Equations of Arbitrary Order," *Ibid.*, **38**, 741–747 (1952); (III) "Assumption of Boundary Values and the Green's Function in the Dirichlet Problem for the General Linear Elliptic Equation," *Ibid.*, **39**, 179–184 (1953).

[2] The existence of a weak (generalized) solution has been shown by Milgram and Lax using the Hille-Yosida method without any assumption of self-adjointness.

[3] Browder, F. E., "Linear Elliptic Systems of Differential Equations," to appear.

[4] Cf. (III), section 2.

[5] Cf. (II).

[6] Cf. Courant-Hilbert, "Methoden der Mathematische Physik," v. 2, Berlin, 1937, Chap. VII.

[7] Cf. Note 6, pp. 511–523.

STOCHASTIC PROCESSES WITH FINITE AND NON-FINITE VARIANCE

BY S. BOCHNER

PRINCETON UNIVERSITY

Communicated December 29, 1952

It has been shown by M. Loève[1] (and perhaps had been anticipated by Kolmogoroff[2]) that if a random point function $x(t)$ on an arbitrary t-set T has finite first and second moments then there exists on T another such function $x^0(t)$ which has the same first and second moments and is Gaussian as well; the latter meaning that for any points $t_1, \ldots, t_k$ in T the joint distribution function of the k random variables $x_j = x^0(t_j)$, $j = 1, \ldots, k$, is a Gaussian one, "regular" or "singular."

We will make two remarks to this theorem. We will extend it to stochastic processes other and more general than are random point functions; and we will also establish it for certain processes for which the second mo-

ments are not finite but in which a Gaussian factor will be isolable nevertheless. And being on the topic of non-finite moments of second order we will in an appendix make some incidental remarks, for ordinary distribution functions, on finite absolute moments of fractional order less than two, and the appendix will be quite independent from the preceding text.

As previously[3] we take a directed set of indices which we will now denote by $\{\lambda\}$ and with each index λ we associate an ensemble

$$\{E_k : (x_j),\ F_k\} \tag{1}$$

in which E_k is a real Euclidean space and F_k is a probability measure on its Borel sets, and we assume as follows. If we take a second index μ, with $\lambda < \mu$, and if the corresponding ensemble is $\{E_l : (y_m),\ F_l\}$ then we have $k \leq l$, and there is given a certain affine transformation

$$x_j = \sum_{m=1}^{l} c_{jm} y_m, \qquad j = 1, \ldots, k, \tag{2}$$

from E_l into all of E_k whose inverse maps the measure F_k into the measure F_l; and if we denote the transformation (2) by $\pi_{\lambda\mu}$ then we have the consistency relation

$$\pi_{\lambda\nu} = \pi_{\lambda\mu}(\pi_{\lambda\nu}) \tag{3}$$

for any $\lambda < \mu < \nu$.

Now, in the present paper, *an (affinely connected) stochastic process* will be any directed family of ensembles (1) with transformations (2) as just described.

It is immediately verified that any real random function $x(t)$ is a process in this sense if each λ is a t-set $(t_1, \ldots, t_k)$ and $\lambda < \mu$ means that the t-set of λ is a subset of the t-set of μ, and F_k is the joint distribution function of the k random variables $x_j = x(t_j)$, $j = 1, \ldots, k$. We note that we then have $c_{jm} = \delta_{jm}$. If we wish to admit complex-valued functions then we can do this by letting $x(t)$ become a random vector of two components. In general, if $x(t)$ is a vector $(x^1(t), \ldots, x^r(t))$ then the Euclidean space corresponding to the index has dimension rk, and the corresponding measure F_{rk} is the joint distribution function of the rk random variables $\{x^\rho(t_j)\}$, $\rho = 1, \ldots, r$; $j = 1, \ldots, k$.

But we may also envisage an additive random set function $X(\Delta)$ as in a previous paper of ours,[4] and in this case an index λ is a partition $(\Delta_1, \ldots, \Delta_k)$, and $\lambda < \mu$ means that μ is a subpartition of it. Also, F_k is the joint distribution function of the random variables $x_j = X(\Delta_j)$, and it so happens that the matrix c_{jm} consists again of numbers $+1$ and 0 only, but the arrangement of these numbers in the matrix is less trivial than before.

THEOREM 1. *If in a stochastic process (1) all distributions F_k have finite (first and) second moments then they can be replaced by normal distributions*

F_k° which have the same first and second moments and are connected by the same affine transformations (2).

Proof: If we introduce the characteristic functions

$$\varphi_\lambda(\alpha) \equiv \varphi_\lambda(\alpha_j) = \int_{E_k} e^{i(\alpha, x)}\, dF_k(x), \tag{4}$$

$$(\alpha, x) = \alpha_1 x_1 + \ldots + \alpha_k x_k,$$

and if for each $\lambda < \mu$ we introduce the transformation

$$\beta_m = \sum_{j=1}^{k} c_{jm}\alpha_j, \qquad m = 1, \ldots, l, \tag{5}$$

which is transposed to (2), and if we denote this transformation by

$$\beta = T_{\mu\lambda}(\alpha) \equiv T(\alpha), \tag{6}$$

then the two characteristic functions $\varphi_\lambda(\alpha_j)$, $\varphi_\mu(\beta_m)$ are connected by the relation

$$\varphi_\lambda(\alpha) \equiv \varphi_\mu(T(\alpha)). \tag{7}$$

Conversely (as we have previously[4] shown for set functions), if we are given a directed family of characteristic functions $\{\varphi_\lambda\}$ and if they are connected by relations (7) for affine transformations (6) with the consistency relations

$$T_{\nu\lambda} = T_{\nu\mu}T_{\mu\lambda},$$

then their distributions constitute a stochastic process as introduced.

It is known that a distribution F_k has finite moments of a given even order r if and only if its characteristic function can be written in the form

$$\varphi_\lambda(\alpha) = P_\lambda^r(\alpha) + o(|\alpha|^r) \tag{8}$$

in the neighborhood of the origin, where $P_\lambda^r(\alpha)$ is an inhomogeneous polynomial of order r. The polynomial in the representation (8) is unique whether r is even or not, and consequently, if for $\lambda < \mu$ we also have

$$\varphi_\mu(\beta) = P_\mu^r(\beta) + o(|\beta|^r)$$

then relation (7) implies the relation

$$P_\lambda^r(\alpha) = P_\mu^r(T(\alpha)). \tag{9}$$

Now, since for a characteristic function we have

$$\varphi(-\alpha) = \overline{\varphi(\alpha)}, \qquad |\varphi(\alpha)| \leq \varphi(0) = 1 \tag{10}$$

it follows immediately that (8) is equivalent with

$$\varphi_\lambda(\alpha) = \exp(Q_\lambda^r(\alpha) + o(|\alpha|^r))$$

where $Q_\lambda^r(\alpha)$ is an inhomogeneous polynomial of order r which is 0 at the origin, and that relation (9) implies relation

$$Q_\lambda^r(\alpha) = Q_\mu^r(T(\alpha)), \tag{11}$$

again for $r \geq 1$. For $r = 1$ this leads to the following very familiar incidental result which we w ll state but loosely.

THEOREM 2. *If in a stochastic process all first moments exist and are finite, then they can be all made to vanish simultaneously by making them zero in the familiar manner in each F_k separately.*

However the interesting case underlying Theorem 1 is $r = 2$, and owing to (10) we than have

$$Q_\lambda^2(\alpha) = i \sum_{p=1}^{k} c_\lambda^p \alpha_p - \sum_{p,\,q=1}^{k} c_\lambda^{pq} \alpha_p \alpha_q, \tag{12}$$

where the first sum is a real linear form and the second sum is a real symmetric positive semidefinite quadratic form. If now we put

$$\varphi_\lambda^0(\alpha) = \exp(Q_\lambda^2(\alpha)),$$

then this by itself is a characteristic function and its distribution $F^\circ(x)$ is Gaussian, and by relation (11) they constitute a stochastic process as described in Theorem 1; *q. e. d.*

THEOREM 3. *If in a stochastic process the second moments are not all finite, but if for every λ we can put*

$$\varphi_\lambda(\alpha) = \exp(\psi_\lambda^1(\alpha) - \psi_\lambda^2(\alpha)), \tag{13}$$

where

$$\psi_\lambda^2(\alpha) = \sum_{p,\,q} c_\lambda^{pq} \alpha_p \alpha_q,$$

and

$$\psi_\lambda^1(\alpha) = i \sum_p c_\lambda^p \alpha_p + \int_{E_k - (0)} \left(e^{i(\alpha, \xi)} - 1 - \frac{i(\alpha, \xi)}{1 + |\xi|^2} \right) d\rho_\lambda(\xi) \tag{14}$$

with

$$d\rho_\lambda(\xi) \geq 0$$

and

$$\int_{E_k - (0)} \frac{|\xi|^2 \, d\rho_\lambda(\xi)}{1 + |\xi|^2} < \infty, \tag{15}$$

then we have separately

$$\psi_\lambda^1(\alpha) = \psi_\mu^1(T(\alpha)), \quad \psi_\lambda^2(\alpha) = \psi_\mu^2(T(\alpha)) \tag{16}$$

and there are two stochastic processes with the distribution functions

$$\{\psi_\lambda{}^1(\alpha)\}, \quad \{\psi_\mu{}^2(\alpha)\}$$

separately.

PROOF: For each λ, $\varphi_\lambda(\alpha)$ is compounded of Gaussian and Poisson addends as in the case of an infinitely subdivisible homogeneous process of the known type, but our affine transformations (6) are very much more general than usual.

Relation (7) implies

$$\psi_\lambda{}^2(\alpha) - \psi_\mu{}^2(T(\alpha)) = \psi_\lambda{}^1(\alpha) - \psi_\mu{}^1(T(\alpha)) \tag{17}$$

and the assertion is that the two functions

$$\psi^m(\alpha) = \psi_\lambda{}^m(\alpha) - \psi_\mu{}^m(T(\alpha)) \qquad m = 1, 2$$

are both identically zero. Now, $\psi^2(\alpha_j)$ is a quadratic form, and if we take any fixed vector $h = (h_1, \ldots, h_k)$ and form the second difference

$$\psi^2(\alpha_j + h_j) - 2\psi^2(\alpha_j) + \psi^2(\alpha_j - h_j)$$

then it has the value

$$2\psi^2(h_j).$$

On the other hand if we form the same difference with $\psi^1(\alpha_j)$ we obtain

$$4\int_{El-(0)} e^{i(T(\alpha),\,\eta)}\, d\sigma_\mu(\eta) - 4\int_{Ek-(0)} e^{i(\alpha,\,\xi)}\, d\sigma_\lambda(\xi), \tag{18}$$

where

$$d\sigma_\mu(\eta) = (\sin(T(h), \eta))^2\, d\rho_\mu(\eta) \geq 0$$

and

$$d\sigma_\lambda(\xi) = (\sin(\alpha, \xi))^2\, d\rho_\lambda(\xi) \geq 0,$$

and also

$$\int_{Ek-(0)} d\sigma_\lambda(\xi) < \infty, \qquad \int_{El-(0)} d\sigma_\mu(\eta) < \infty$$

because of (15). Next, since we have

$$(T(\alpha), \eta) = \sum_{j=1}^{k} \alpha_j \sum_m c_{jm}\alpha_m,$$

it is not hard to conclude that it is possible to find a set function $\sigma_\lambda'(\xi)$ in $E_k - (0)$ which the inverse of (2) maps into $\sigma_\mu(\eta)$ such that

$$\int_{El-(0)} e^{i(T(\alpha),\,\eta)}\, d\sigma_\mu(\eta) = \int_{Ek-(0)} e^{i(\alpha,\,\xi)}\, d\sigma_\lambda'(\xi).$$

Finally if we put

$$d\sigma(\xi) = 2d\sigma_\lambda'(\xi) - 2d\sigma_\lambda(\xi)$$

then we obtain altogether

$$\psi^2(h_j) = \int_{E_k-(0)} e^{i(\alpha,\xi)}\, d\sigma(\xi) \tag{19}$$

where

$$\int_{E_k-(0)} |d\sigma(\xi)| < \infty.$$

We recall that in the preceding derivation $\{h_j\}$ was fixed so that $\psi^2(h_j)$ is a constant in (α_j), and $\sigma(\xi)$ may depend on h, and that (α_j) is variable and varies over $E_k:(\alpha_j)$. Therefore the right side in (19) is a Fourier Stieltjes transform of a set function of bounded variation which has value zero on the point (0). The left side, however, is a constant, and because of the uniqueness theorem for such transforms both sides must be zero, $\psi^2(h_j) = 0$. If we now vary h_j we obtain $\psi^2(\alpha) = 0$, and therefore also $\psi^1(\alpha) = 0$ by (17), which completes the proof of the theorem.

Appendix. If

$$\varphi(\alpha) = \int_{-\infty}^{\infty} e^{i\alpha x} dF(x) \tag{20}$$

is a characteristic function in E_1, then we have

$$1 - \frac{\varphi(\alpha) + \varphi(-\alpha)}{2} = 2\int_{-\infty}^{\infty} \left(\sin\frac{\alpha}{2}x\right)^2 dF(x),$$

and hence for any function $\lambda(\alpha) > 0$ in $0 < \alpha \leq 1$ we have

$$\int_{+0}^{1} \frac{1}{\lambda(\alpha)} \left(1 - \frac{\varphi(\alpha) + \varphi(-\alpha)}{2}\right) d\alpha = \int_{-\infty}^{\infty} \mu(x) dF(x) \tag{21}$$

where in $0 \leq x < \infty$ we have

$$\mu(-x) = \mu(x) = 2\int_0^1 \frac{\left(\sin\frac{\alpha}{2}x\right)^2}{\lambda(\alpha)}\, d\alpha. \tag{22}$$

Also, the relation (21) is valid whether the integrals are finite or $+\infty$, provided only $\mu(x)$ is finite and continuous, and for this it is sufficient that

$$\int_0^1 \frac{\alpha^2}{\lambda(\alpha)}\, d\alpha < \infty. \tag{23}$$

In the leading case $\lambda(\alpha) = \alpha^{q+1}$, $0 < q < 2$, we have

$$\mu(x) = 2\int_0^1 \frac{\left(\sin\frac{\alpha}{2}x\right)^2}{\alpha^{q+1}}\, d\alpha = x^q \cdot 2\int_0^x \frac{\left(\sin\frac{\alpha}{2}\right)^2}{\alpha^{q+1}}\, d\alpha,$$

and it is not hard to obtain the following conclusion.

THEOREM 4. *For a characteristic function (20) the limit*

$$\lim_{\epsilon \to 0} \int_{\epsilon}^{1} \left(1 - \frac{\varphi(\alpha) + \varphi(-\alpha)}{2}\right) \frac{d\alpha}{\alpha^{q+1}}$$

exists and is finite, $0 < q < 2$, if and only if we have

$$\int_{-\infty}^{\infty} |x|^q \, dF(x) < \infty.$$

For instance, if we consider for $0 < p < 2$ the stable characteristic function

$$e^{-|\alpha|^p} = \int_{-\infty}^{\infty} e^{i\alpha x} f_p(x) \, dx,$$

then we have

$$\int_{-\infty}^{\infty} |x|^q f_p(x) \, dx \begin{cases} < \infty, & q \leq p \\ = \infty, & q \geq p \end{cases}$$

and our present proof of Theorem 4 is probably the quickest way for proving this. However the result can also be refined, and if, for instance, we put $\lambda(\alpha) = \alpha^{p+1} |\log \alpha|^{\sigma}$ then we obtain from (21) the logarithmic refinement

$$\int_{-\infty}^{\infty} |x|^p \, |\log x|^{\sigma} f_p(x) \, dx \begin{cases} < \infty, & \sigma < -1 \\ = \infty, & \sigma \geq -1. \end{cases}$$

Finally, if we wish to extend relation (21) to a multidimensional characteristic function

$$\varphi(\alpha_j) = \int_{E_k} e^{i(\alpha, x)} \, dF_k(x),$$

then we first of all require a suitable analog to the arithmetic mean

$$\psi(|\alpha|) = {}^1/_2(\varphi(\alpha) + \varphi(-\alpha)),$$

itself a characteristic function. It turns out that it is the spherical mean over spheres with center (0),

$$\psi(|\alpha|) = \frac{1}{\omega_k} \int_{S_{k-1}} \varphi(|\alpha|\xi_1, \ldots, |\alpha|\xi_k) \, d\omega_\xi, \tag{24}$$

$\xi_1^2 + \ldots + \xi_k^2 = 1$, itself a characteristic function, which can also be written as

$$\psi(|\alpha|) = \int_{E_k} \Pi_k(|\alpha| \cdot |x|) \, dF_k(x_j)$$

where $\Pi_k(t) = V_l(t)/V_l(0)$ and $V_l(t) = J_l(t) \cdot t^{-l}$, $l = {}^1/_2(k - 1)$, with $J_l(t)$ being the Bessel function.

The analog to (21) is then

$$\int_{+0}^{1} (1 - \psi(\beta)) \frac{d\beta}{\lambda(\beta)} = \int_{E_k} \mu(x_j) dF_k(x_j)$$

where

$$\mu(x_j) = \int_0^1 (1 - H_k(|\beta| \cdot |x|)) \frac{d\beta}{\lambda(\beta)}.$$

Now, for $k \geq 2$ we have $1 - H_k(t) = 0(t^2)$ as $t \to 0$, just as for $k = 1$, but as $t \to \infty$ we have $H_k(t) = o(1)$ for $k \geq 2$, whereas for $k = 1$ it is only $O(1)$, and the following theorem ensues.

THEOREM 5. *For $k \geq 2$, if $\lambda(\alpha) > 0$ in $0 < \alpha \leq 1$, and (23) holds, if $\varphi(\alpha_j)$ is any characteristic function in E_k and $\psi(|\alpha|)$ is its spherical mean (24), then we have*

$$\int_{+0}^1 \frac{1 - \psi(\alpha)}{\lambda(\alpha)} d\alpha < + \infty$$

if and only if

$$\int_{|x| \geq 1} \left(\int_{|x|^{-1}}^1 \frac{d\alpha}{\lambda(\alpha)} \right) dF_k(x_j) \infty.$$

[1] In his appendix to: Levy, P., *Processus Stochastiques*, Paris (1948).

[2] See the bottom of page 609 in: Khinchin, A., Korrelations theorie der stationären stochastischen Prozesse, *Math. Annalen*, **109**, 604–615 (1933).

[3] Bochner, S., "Partial Ordering in Theory of Stochastic Processes," these PROCEEDINGS, **36**, 439–443 (1950).

[4] Bochner, S., "Stochastic Processes," *Ann. Math.*, **48**, 1014–1061 (1947).

---

*A PROPERTY OF MINIMAL SURFACES**

BY ROBERT FINN

INSTITUTE FOR ADVANCED STUDY

Communicated by Marston Morse, January 14, 1953

1. It is well known that if $f(x, y)$ is harmonic in an open set D of the plane and bounded in magnitude by a constant M, then at every point $X_0 = (x_0, y_0)$ of D there is a bound on $f_x^2 + f_y^2$ which depends on M and the distance of (x_0, y_0) to the boundary of D, but not on the particular function considered. Since it is intuitively clear that minimal surfaces enjoy exceptional smoothness properties, it is natural to expect that a corresponding theorem is valid for minimal surfaces which can be represented by a single valued function $\varphi(x, y)$. In this note we state a theorem of this type which is valid for solutions of the minimal surface equation and for solutions of a related class of equations.

Except where otherwise noted, we shall assume that all functions considered are twice continuously differentiable.

THEOREM 1. *Let D be a bounded open set of the plane, let d be a positive number; let F be the set of all points interior to D and of distance not less than d from the boundary Γ of D. Then there exists a positive number $K_D(\mathfrak{A}, d)$ with the following properties:*

(i) if $\varphi(x, y)$ is a single-valued solution in D of the minimal surface equation

$$(1 + q^2)\varphi_{xx} - 2pq\varphi_{xy} + (1 + p^2)\varphi_{yy} = 0,\ p = \varphi_x,\ q = \varphi_y, \quad (1)$$

and if the area of the surface represented by $\varphi(x, y)$ is smaller than a finite number $\mathfrak{A}$, then $\varphi_x^2 + \varphi_y^2 \leq K_D(\mathfrak{A}, d)$ at any point of F.

ii) $K_D(\mathfrak{A}, d)$ is non-decreasing in $\mathfrak{A}$ for every fixed d, non-increasing in d for every fixed $\mathfrak{A}$.

The following corollary is a simple consequence of Theorem 1.

COROLLARY. *There exists a function $K_D^*(M, d)$ such that if $\varphi(x, y)$ is a solution of (1) in D and if $|\varphi(x, y)| < M < \infty$ then $\varphi_x^2 + \varphi_y^2 \leq K_D^*(M, d)$ at any point of F.*[1]

For F may be covered by a finite number N_F of circles of radius $d/4$. To each of these circles we associate a circle with the same center, but of radius $d/2$. This defines a new family of circles which still covers F and lies interior to D. Let Δ be the union of all circles of the family, Γ_Δ the boundary of Δ. Then each point of F is of distance not less than $d/4$ from Γ_Δ. Denote one of these circles by C, its boundary by Γ_C. Then if $\varphi(x, y)$ is a solution of (1) in D for which $|\varphi(x, y)| < M$,

$$-\pi d^2/4 + \int\int_C \sqrt{1 + p^2 + q^2}\, dxdy \leq \int\int_C \frac{p^2 + q^2}{\sqrt{1 + p^2 + q^2}}\, dxdy = \oint_{\Gamma_C} \frac{\varphi}{\sqrt{1 + p^2 + q^2}} \frac{d\varphi}{dn}\, ds \leq M \oint_{\Gamma_C} \frac{\sqrt{p^2 + q^2}}{\sqrt{1 + p^2 + q^2}}\, ds \leq \pi d M.$$

We thus have for the area $\mathfrak{A}_\Delta$ of that part of the surface represented by $\varphi(x, y)$ which lies above Δ,

$$\mathfrak{A}_\Delta \leq \pi d N_F \cdot (M + d/4),$$

from which we find, using Theorem 1, that at any point of F,

$$\varphi_x^2 + \varphi_y^2 \leq K_\Delta(\mathfrak{A}_\Delta, d/4) \leq K_\Delta[\pi d N_F \cdot (M + d/4), d/4].$$

We include here an outline of the proof of Theorem 1. The property (ii) is a consequence of property (i), since for each pair $(\mathfrak{A}, d)$ we may choose for $K_D(\mathfrak{A}, d)$ the lower bound of all numbers for which (i) is satisfied. If for some $(\mathfrak{A}, d)$ there did not exist a finite number with property (i) then there would be a sequence of solutions $\{\varphi^{(n)}(x, y)\}$ of (1) in D representing

surfaces uniformly bounded in area, and a sequence of points $\{(x_n, y_n)\} = \{X_n\}$ converging to a point $(x_0, y_0) = X_0$ in F, such that $\varphi_x^{(n)2}(x_n, y_n) + \varphi_y^{(n)2}(x_n, y_n) \to \infty$ as $n \to \infty$. Let C_0 be a circle about X_0 of radius $d/2$. For each n, it is possible to introduce isothermal parameters $[\xi^{(n)}(x, y), \eta^{(n)}(x, y)]$ on the surface $\varphi^{(n)}(x, y)$ which map the closure $\bar{C}_0$ of C_0 in a 1–1 manner on the closure $\bar{C}_0^*$ of a circle C_0^*: $\xi^2 + \eta^2 < 1$. The mapping may be performed in such a manner that three given points on the boundary of C_0 correspond to three given points on the boundary of C_0^*, for every n. A slight modification of a technique due to Lebesgue[2] then shows that it is possible to select a subsequence of the vectors $\{x^{(n)}(\xi, \eta) + iy^{(n)}(\xi, \eta)\}$ which converges uniformly on $\bar{C}_0^*$ to a continuous mapping onto $\bar{C}_0$. On the other hand, it is easily verified that the function

$$w^{(n)}(\xi + i\eta) = \frac{\varphi_x^{(n)}[x^{(n)}(\xi, \eta), y^{(n)}(\xi, \eta)] - i\varphi_y^{(n)}[x^{(n)}(\xi, \eta), y^{(n)}(\xi, \eta)]}{1 + \sqrt{1 + \varphi_x^{(n)2} + \varphi_y^{(n)2}}}$$

is an analytic function of $\xi + i\eta$, and $|w^{(n)}| < 1$ for all n. Thus there is a subsequence of the $\{w^{(n)}\}$ which converges uniformly on every concentric subcircle of C_0^* to an analytic function $w^*(\xi + i\eta)$, and either $|w^*| < 1$ in C_0^* or $w^* \equiv e^{i\alpha}$ for some real constant α. The former case leads to a contradiction with the properties of the sequence of points $\{X_n\}$ introduced above, the latter case would be in contradiction with the hypothesis of bounded area.

2. The properties of minimal surfaces described above admit an immediate generalization. *There exists a well-defined class* $\mathfrak{C}$ *of quasi-linear elliptic partial differential equations, similar in structure to (1), for which the conclusions of Theorem 1 and of its corollary are valid.* The class $\mathfrak{C}$ includes in particular all equations of the form

$$(\rho\varphi_x)_x + (\rho\varphi_y)_y = 0, \tag{2}$$

$$\rho = 1/W + \lambda\epsilon(t),\ W^2 = 1 + t,\ t = \varphi_x^2 + \varphi_y^2,$$

provided that $\epsilon(t) = 0(t^{-3/2})$, $\left|\frac{d\epsilon}{dt}\right| = 0(t^{-5/2})$ as $t \to \infty$ and λ is a real number sufficiently small that $\rho W^4 \cdot [\rho(t) + 2\rho'(t) \cdot t] > h^2 > 0$ for all non-negative t. The proof involves a discussion of the relation between two Riemannian manifolds, each of which attaches in a natural way to every solution of the equation. The following lemma is applied.

LEMMA. *Let* $\{w^{(n)}\}$ *be a sequence of complex valued functions, each twice continuously differentiable in a plane domain* D. *We assume that* $|w^{(n)}| < 1$ *and that the* $\{w^{(n)}\}$ *are of uniformly bounded eccentricity*[3] *for all* n. *Then there exists a subsequence which converges uniformly on every compact sub-*

domain F of D. The limit function w^* satisfies the maximum modulus principle in the strict sense, that is, if $M = \sup_D |w^*|$ then either $|w^*| < M$ at every point of D, or $w^* \equiv Me^{i\alpha}$ for some real constant α.

In fact, it can be shown that either w^* is interior in the sense of Stoilow, or $w^* \equiv$ constant. A proof of the lemma follows very readily from results of Morrey[4] and from the technique of Lebesgue already cited.

3. Let $\varphi(x, y)$ be a solution in an open set D of the quasi-linear elliptic equation

$$a(x, y, \varphi_x, \varphi_y)\varphi_{xx} + 2b(x, y, \varphi_x, \varphi_y)\varphi_{xy} + c(x, y, \varphi_x, \varphi_y)\varphi_{yy} = 0, \qquad (3)$$
$$ac - b^2 = 1,$$

and let C be a circle of radius d which lies entirely in D. Then it is known[5] that in the concentric subcircle of radius $d/2$, the second derivatives of $\varphi(x, y)$ satisfy a Hölder condition which depends only on d and on $\sup_C (\varphi_x^2 + \varphi_y^2)$. We thus obtain

THEOREM 2. *Let $\{\varphi^{(n)}(x, y)\}$ be a sequence of single-valued solutions in an open connected set D of an equation of class $\mathfrak{C}$, and suppose that the surfaces represented by these functions are uniformly bounded in area for all n. Suppose further that there is a point (x_0, y_0) of D at which $|\varphi^{(n)}(x_0, y_0)| \leq K < \infty$ for every n. Then there is a subsequence of solutions which, together with the corresponding sequences of derivatives up to the second order, converges throughout D. The convergence is uniform on every compact subdomain of D and the limit function is a solution of the equation in D.*

A corresponding theorem is valid for a sequence of solutions which are uniformly bounded in magnitude in D.

4. Let D be a simply connected domain, Γ its boundary. Let φ^* be a function defined on Γ. Given an equation of type (3), the natural analogue of the Dirichlet problem is to seek a solution in D which assumes the values φ^* on Γ. General existence theorems have been obtained, first by S. Bernstein,[6] later by Leray and Schauder[7] and by others, under the assumptions that D is convex and that the boundary curve Γ^* determined by φ^* satisfies a three point condition.[8] For linear equations neither of these conditions is necessary. On the other hand, it is known that there exists a boundary curve Γ^* which projects into a Jordan curve in the (x, y) plane, and which is such that, (i) there is a unique minimal surface which spans Γ^*, (ii) this surface has a multiply covered (x, y) projection. Thus it is not in general possible to eliminate the restriction on D. It is an open question, however, whether this is equally true for the three-point condition. We are now prepared to give a partial answer to this question.

THEOREM 3. *Let D be a convex domain, Γ its boundary. We assume that Γ can be described by twice continuously differentiable functions of arc length, and that the curvature of Γ does not vanish. Let φ^* be a continous function defined on Γ. Then for any equation of class $\mathfrak{C}$, there exists a unique func-*

tion $\varphi(x, y)$ which is a solution throughout D and which attaches continuously to φ^* on Γ.

It is a straightforward matter to show that if a sequence of solutions of a quasi-linear elliptic equation converges uniformly on the boundary of a domain, then it converges uniformly throughout the closure of the domain. But if Γ has non-vanishing curvature, any function continuous on Γ can be approximated uniformly by functions which yield boundary curves satisfying three-point conditions.

* This investigation was supported by the Office of Naval Research.

[1] The author has knowledge of an unpublished proof of this property of minimal surfaces, based on other methods, due to Professor M. Shiffman.

[2] Lebesgue, H., *Rend. del Circ. Mat. di Palermo*, **24,** 386–388 (1907).

[3] The eccentricity e of $w = u + iv$ at a point $z = x + iy$ is defined as $\frac{1}{2}\left(\frac{Q_1}{Q_2} + \frac{Q_2}{Q_1}\right)$, where $Q_1 = \overline{\lim}\left|\frac{\Delta w}{\Delta z}\right|$, $Q_2 = \underline{\lim}\left|\frac{\Delta w}{\Delta z}\right|$ as $\Delta z \to 0$. If u and v have continuous first derivatives, then $e = \dfrac{u_x^2 + u_y^2 + v_x^2 + v_y^2}{2J\binom{uv}{xy}}$ wherever the Jacobian does not vanish.

[4] Morrey, C. B., Jr., *Trans. Am. Math. Soc.*, **43,** 141–145 (1938).

[5] Morrey, C. B., Jr., *loc. cit.*, pp. 153–158, and Hopf, E., *Math. Zeit.*, **30,** 404–413 (1929).

[6] Bernstein, S., *Math. Ann.*, **69,** 126 (1910).

[7] Leray, J., and Schauder, J., *Ann. École Normale Sup.*, **51,** 69–70 (1934).

[8] A boundary curve is said to satisfy a three-point condition with constant Δ provided that any plane which intersects the curve in three or more points has maximum inclination less than Δ. It is known that Δ serves as an *a priori* bound on the gradient of any solution of (3) in D which assumes the values φ^* on Γ. For a detailed discussion of conditions under which the three-point condition is satisfied, see Schauder, J., *Math. Zeit.*, **37,** 625–628 (1933).

SECOND ORDER ELLIPTIC SYSTEMS OF DIFFERENTIAL EQUATIONS

BY CHARLES B. MORREY, JR.

UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated by S. Bochner, January 7, 1953

In this paper, we consider the existence and regularity properties of the solutions of linear elliptic systems of partial differential equations in N dependent and ν independent variables $x = (x^1, \ldots, x^\nu)$ of the form

$$\int_{R^*}(a_{ij}^{\alpha\beta}u_\beta^j + b_{ij}^\alpha u^j + e_i^\alpha)\,dx_\alpha' = \int_R (c_{ij}^\alpha u_\alpha^j + d_{ij}\,u^j + f_i)\,dx,$$
$$i, j = 1, \ldots, N;\ \alpha, \beta = 1, \ldots, \nu;\ dx_\alpha' = (\zeta_\alpha \cdot n)\,dS; \quad (1)$$

in which ζ_α is the unit vector in the x^α direction and n is the exterior normal. In (1) and throughout this note we have used the usual tensor summation convention; Greek indices run from 1 to ν and Latin indices from 1 to N. The coefficients $a(x)$ are assumed to be continuous on the closure G of the bounded domain G, $b(x)$, $c(x)$, and $d(x)$ are assumed to be bounded and measurable, and $e(x)$ and $f(x)$ to be in L_2 on G. The solutions $u = (u^1, \ldots, u^N)$ are required to be of class $\mathfrak{B}_2{}^{1,2}$ on domains interior to G and to satisfy (1) on "almost all cells".[1] The ellipticity condition is that the $N \times N$ determinant

$$\left| a_{ij}^{\alpha\beta}(x)\lambda_\alpha\lambda_\beta \right| \neq 0, \ \lambda \neq 0, \ x \in G \ (\lambda = \lambda_1, \ldots, \lambda_\nu, \text{ real}). \qquad (2)$$

We have also considered second order systems of the form

$$a_{ij}^{\alpha\beta}u^j_{x_\alpha x_\beta} = c_{ij}^{\alpha}u^j_{x_\alpha} + d_{ij}u^j + f_i \qquad (3)$$

in which the solutions u and their first generalized derivatives are required to be of class $\mathfrak{B}_2$ on domains interior to G and u is required to satisfy (3) almost everywhere.

The results for the general elliptic system (3) which hold on domains interior to G are as follows:

1. If a solution u of (3) is in L_2 on G and $\bar{G}_1 \subset G$, then

$$d_2(u, G_1) = \left\{ \int_{G_1} \sum_{i, \alpha, \beta} (u^i_{x_\alpha x_\beta})^2 \, dx \right\}^{1/2} \leqq K_1$$

where K_1 depends only on ν, N, bounds for the coefficients c, and d, bounds for the norms in L_2 of u and f, G_1 and its distance from G^*, the modulus of continuity of the tensor $a(x)$ on $\bar{G}$, and a certain "ellipticity bound" M for $a(x)$, M being $\geqq 1$, the equality holding for Laplace's equation.

2. If, in addition to the hypotheses above, we assume that f satisfies an integral growth condition of the form

$$L_2[f, C(P, r)] = \int_{C(P, r)} \left(\sum_{i=1}^{N} f_i^2 \right) dx \leqq L(r/\delta_P)^{\nu - 2 + 2\mu} \qquad 0 < \mu < 1, \quad 0 \leqq r \leqq \delta_P. \quad (4)$$

on every sphere $C(P, r)$ with center at P and radius r, δ_P being the distance of P from the boundary G^*, then u is of class C'_μ on G, i.e., u is of class C' on G and its derivatives satisfy uniform Hölder conditions with exponent μ on each $G_1 \subset G$ in which the Hölder coefficient depends only on the quantities in 1. and on μ and L.

3. If, also, all the coefficients $a(x)$, $c(x)$, $d(x)$, and $f(x)$ are of class $C_\mu^{(n)}$ on $\bar{G}$, $0 < \mu < 1$, $n \geqq 0$, then u is of class $C_\mu^{(n+2)}$ on G, the bounds and Hölder coefficients for the derivatives of u on a $G_1 \subset G$ depending only on the quantities in 1. and 2. and bounds and Hölder coefficients for $a(x)$ to $f(x)$ on $\bar{G}$.

For the system (1) we have obtained the following results:

4. If a solution u of (1) is in L_2 on G and $\bar{G} \subset G$, then

$$d_1(u, G_1) = \left\{ \int_{G_1} \sum_{i,\alpha} (u^i_{x^\alpha})^2 \, dx \right\}^{1/2} \leqq K_2$$

where K_2 depends on the quantities in 1. and bounds for $b(x)$ and the norm of e in L_2 on G.

5. If, also, e and f satisfy the integral growth condition (4), then the first derivatives of u satisfy similar growth conditions and u is of class $C^{(0)}_\mu$ on G, the Hölder coefficient for u on any $\bar{G}_1 \subset G$ depending only on the quantities in 4. and on u and L.

6. If, also, $a(x)$, $b(x)$, and $e(x)$ are of class $C^{(0)}_\mu$, $0 < u < 1$, on G, then u is of class C'_μ on G, the Hölder coefficients for the first derivatives of u on any $\bar{G}_1 \subset G$ depending only on those for $a(x)$, $b(x)$, and $e(x)$ on $\bar{G}$ and the quantities in 5.

7. If, also, $a(x)$, $b(x)$, and $e(x)$ are of class $C^{(n+1)}_\mu$ and $c(x)$, $d(x)$, and $f(x)$ are of class $C^{(n)}_\mu$ on $\bar{G}$, $n \geqq 0$, $0 < \mu < 1$, then u is of class $C^{(n+2)}_\mu$ on G, the bounds and Hölder coefficients for u and its derivatives on any $\bar{G}_1 \subset G$ depending only on these bounds for $a(x) - f(x)$ and on the other quanities in 6.

Of course, it is necessary to prove 7. only for the case $n = 0$ when it follows that u, being of class C^*_μ, satisfies

$$\frac{\partial}{\partial x^\alpha} (a^{\alpha\beta}_{ij} u^j_{x^\beta} + b^\alpha_{ij} u^j + e^\alpha_i) = c^\alpha_{ij} u^j_{x^\alpha} + d_{ij} u^j + f_i$$

which reduces to a system (3).

In order to obtain existence theorems and results concerning regularity on the boundaries of regular domains, the writer found it necessary to discuss only the system (1) and to confine himself to *strongly elliptic systems* which satisfy

$$a^{\alpha\beta}_{ij}(x) \lambda_\alpha \lambda_\beta \xi^i \xi^j > 0, \; x \in \bar{G}, \; \lambda \neq 0, \; \xi \neq 0.$$

Some such restriction is probably necessary since simple examples[3] show that the Dirichlet problem for a general elliptic system with constant coefficients and only the a's present is not always uniquely solvable. Such an example is the system

$$\begin{aligned} (u^1_{x^1x^1} + 2u^1_{x^1x^2} - u^1_{x^2x^2}) + (u^2_{x^1x^1} - 2u^2_{x^1x^2} - u^2_{x^2x^2}) &= 0 \\ -(u^1_{x^1x^1} - 2u^1_{x^1x^2} - u^1_{x^2x^2}) + (u^2_{x^1x^1} + 2u^2_{x^1x^2} - u^2_{x^2x^2}) &= 0 \end{aligned}$$

for which the determinant (2) equals $2(\lambda_1^2 + \lambda_2^2)^2$ but which has the solution $u^1 = u^2 = 1 - (x^1)^2 - (x^2)^2$. On the other hand it is easy to prove that any general system satisfying the *symmetry conditions*

$$a^{\alpha\beta}_{ij}(x) + a^{\beta\alpha}_{ij}(x) = a^{\alpha\beta}_{ji}(x) + a^{\beta\alpha}_{ji}(x) \qquad (\text{all } \alpha, \beta, i, j)$$

can be reduced to a strongly elliptic system by merely choosing proper linear combinations of the given equations with *constant coefficients*.

For strongly elliptic systems, we have shown that *the results stated in 4.–7. hold for domains G_1 which intersect the boundary G^* of G provided that the part of G^* in G_1 is sufficiently regular, u vanishes on that part of G^*, and e and f satisfy their previous hypotheses on G_1; in 4. and 5. the given part of G^* is required to be only of class C^1, in 6. it is required to be of class C_μ^1, and in 7. to be of class $C_\mu^{(n+2)}$. The cases where u does not vanish on the given part of G^* may be handled by subtracting off an appropriate function coinciding with u there: if, in 6. and 7., the boundary values of u are of class C_μ^1 or $C_\mu^{(n+2)}$ along G^*, we may subtract off a harmonic function, otherwise u must coincide on $G^* \cap G_1$, with some function u^* which has the regularity properties which hold for u when $u = 0$ on G^*.*

For the strong elliptic system (1) with d_{ij} replaced by $d_{ij} + \lambda\delta_{ij}$, δ_{ij} being the Kronecker delta (this is analogous to considering an equation $L(u) = \lambda u$, L being a differential operator), we have shown that *if G is an arbitrary bounded domain, the λ-system* (1) (with the most general coefficients allowed) *has a unique solution u which coincides (in $\mathfrak{P}_2$) on G^* with any given function u^* in $\mathfrak{P}_2$ on G, provided λ is not in an isolated set of characteristic values which is bounded above; if λ is characteristic, the homogeneous system ($e = f = 0$) has a finite dimensional manifold of non-zero solutions which vanish on G^*.* The proof of these results involves operators in real Hilbert space as do the recent proofs of Browder[4] for the single equation of order $2m$ and of Višik[5] for the strongly elliptic system; some modification of their methods is needed to handle the case of our general coefficients.

Finally, the results above and the methods used to obtain them are applied to obtain results concerning non-linear systems as follows:

8. If z is of class C'' and is the solution on G of a general elliptic second order system

$$\varphi_i(x, z, z_{x^\alpha}, z_{x^\alpha x^\beta}) = 0$$

in which the φ_i are of class C', then z is of class C_μ'' on G for any μ, $0 < \mu < 1$, and its second derivatives are of class $\mathfrak{P}_2$ on any interior domain; if, also, the φ_i are of class $C_\mu^{(n)}$, $n \geqq 1$, $0 < \mu < 1$, then z is of class $C_\mu^{(n+2)}$ on G.

9. If z is of class C' and is the solution of a regular variational problem, i.e., is the solution of Haar's equations

$$\int_{R^*} f_{p\alpha^i}\, dx_\alpha^i = \int_R f_{z_i}\, dx,$$

in which f and its derivatives with respect to the p_α^i are of class C', then z is of class C_μ^1 on G for any μ, $0 < \mu < 1$, and the first derivatives of z are of class $\mathfrak{P}_2$ on any interior domain. If, also, f and the $f_{p\alpha^i}$ are of class $C_\mu^{(n)}$ $n \geqq 1$, then z is of class $C_\mu^{(n+1)}$ on G. Boundary value results analogous to those for systems (1) also hold.

The method of proof consists in first studying the general elliptic system (1) in the special case where the coefficients $a(x)$ are constants and all the

other functions except $e(x)$ vanish. The method of F. John[6] is easily adapted to show the existence of an elementary matrix $\Gamma^{ij}(y)$ for the resulting homogeneous system. The $\Gamma^{ij}(y)$ are seen to be analytic if $y \neq 0$, positively homogeneous of degree $2 - \nu$ in y if $\nu > 2$ and of the form

$$C^{ij} \log |y| + \tilde{\Gamma}^{ij}(y)$$

if $\nu = 2$, the C^{ij} being constants and the $\tilde{\Gamma}^{ij}$ being positively homogeneous of degree 0; in all cases $\Gamma^{ij}(-y) = \Gamma^{ij}(y)$. Using this elementary matrix, the analyticity and certain boundedness properties on interior domains of "$\bar{a}$-harmonic" vectors (i.e., solutions of the homogeneous system) are established, and it is shown that a vector which is $\bar{a}$-harmonic with finite Dirichlet integral on the whole space is a constant. If e is in L_2 on some sphere $C(P, a)$, special solutions of the non-homogeneous system are given by

$$u^j(x) = \int_{C(P,a)} \Gamma^{jk}_{\beta}(\xi - x) e^{\beta}_{k}(\xi)\, d\xi, \quad \Gamma_{\beta}(y) = \Gamma_{,\beta}(y) \tag{5}$$

This is seen to be true if e is of class C'' everywhere and vanishes outside $C(P, a)$. In this case the first and second derivatives of u are seen to be continuous and in L_2 over the whole space. By taking Fourier transforms, one finds that if u is given by (5) then

$$D_2(u, E_\nu) \leqq M\, L_2[e, C(P, a)], \tag{6}$$

M being the "ellipticity bound" mentioned in 1. The results (5) and (6) follow for the general e in L_2 by approximations. A tricky "Dirichlet growth" theorem shows that if e satisfies an integral growth condition (4) at P with any exponent $2\lambda < \nu$, then the first derivatives of u (given by (5)) satisfy a similar condition; next, if e satisfies a certain type of Hölder condition on $C(P, a)$, a classical proof shows that the first derivatives of u satisfy similar conditions; bounds are obtained in all cases. Finally if f is in L_2 on $C(P, a)$, its $\bar{a}$-potential is defined in the obvious way; the properties of these follow from the fact that their first derivatives are of the form (5).

The results 1–3 ($n = 0$) and 4–6 for the general systems (1) and (3) are reduced to those for the simplified systems above by confining ourselves to sufficiently small neighborhoods $C(P, a)$ of each interior point P and introducing certain transformations from u to U satisfying

$$\bar{a}^{\alpha\beta}_{ij} U^{j}_{x^\alpha x^\beta} + (a^{\alpha\beta}_{ij} - \bar{a}^{\alpha\beta}_{ij}) u^{j}_{x^\alpha x^\beta} = c^{\alpha}_{ij} u^{j}_{x^\alpha} + d_j u^i \qquad (\bar{a} = a(P))$$

in the case of (3), for instance. This idea has to be combined with the methods of Schauder[7] in order to obtain the bounds in 1. and 4. The higher regularity properties in 3. and 7. are obtained by using the device of Lichtenstein[8, 9] on interior domains and the uniform boundedness results in 1–6. This device is also the tool used in obtaining the results in 8. and 9.

In the case of the strongly elliptic system we use a theorem first proved in a special case by Shiffman[10] and an integral theorem proved by Van Hove[11] to establish the existence and uniqueness theorems for the solutions of the simplified systems on arbitrary bounded domains, sharp bounds for Dirichlet integrals, etc. These results are used to show that if u is a-harmonic and in $\mathfrak{B}_2$ on a hemisphere and vanishes (in the $\mathfrak{B}_2$ sense) on the flat boundary, then u is analytic on that flat boundary; bounds are obtained, etc. Then the results previously obtained for spheres are now extended to parts of spheres ($\geqq$ hemispheres) on one side of a hyperplane and the continuity properties are shown to hold on the flat part of the boundary for functions u which vanish on that part (some quite new and rather surprising "Dirichlet growth" theorems occur in this development). The results for the general system (1) on such domains are obtained by introducing operators as before. The higher order regularity properties are first proved by a modified device of Lichtenstein to hold up to the flat boundaries of these special domains and are then obtained for sufficiently small boundary neighborhoods of (locally) regular domains G by mapping these onto the special domains.

[1] Calkin, J. W., "Functions of Several Variables and Absolute Continuity, I," *Duke Math. J.*, **6**, 170–186 (1940).

[2] Morrey, C. B., Jr., "Functions of Several Variables and Absolute Continuity, II," *Ibid.*, **6**, 187–215 (1940).

[3] Bicadze, A. V., "On the Uniqueness of the Solution of the Dirichlet Problem for Elliptic Partial Differential Equations," *Uspehi Matem. Nauk (N.S.)*, **3**, 211–212 (1948).

[4] Browder, F. E., "The Dirichlet Problem for Linear Elliptic Equations of Arbitrary Even Order with Variable Coefficients," PROC. NATL. ACAD. SCI., **38**, 230–235 (1952).

[5] Višic, M. I., "On Strongly Elliptic Systems of Differential Equations," *Doklady Acad. Nauk S.S.S.R. (N.S.)*, **74**, 881–884 (1950). Russian.

[6] John F., "The Fundamental Solution of Linear Elliptic Differential Equations with Analytic Coefficients," *Comm. Pure Appl. Math.*, **3**, 273–304 (1950).

[7] Schauder, J., "Ueber lineare elliptische differentialgleichungen zweiter ordnung," *Math. Zeit.*, **38**, 257–282 (1934).

[8] Lichtenstein, L., "Ueber den analytischen charakter der lösungen zweidimensionaler variationsprobleme," *Bull. l'Acad. Sci. Cracovie, Class. Sci. Math. Nat.*, Ser. A, **1912**, 915–941.

[9] Morrey, C. B., Jr., "Multiple Integral Problems in the Calculus of Variations and Related Topics" (Chapter VII), *Univ. Calif. Pub. Math, New Ser.*, **1**, 1–130 (1943).

[10] Shiffman, Max, "Differentiability and analyticity of solutions of double integral variational problems." *Annals of Math.* **48**, pp. 274–284 (1947) (theorem 3).

[11] Van Hove, L., "Sur l'extension de la condition de Legendre du Calcul des Variations aux integrales multiples à plusieurs fonctions inconnues," *Nederl. Akad. Wetensch.* **50**, pp. 18–23 (1947).

SPIRALS IN THE PLANE

BY R. L. MOORE

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated December 22, 1952

In the plane, let J denote a definite circle, let O denote a definite point of J, let I denote the interior of J, let A denote a point of I and let α denote the straight line interval whose end-points are O and A. Let Z denote a definite point of α and LZ a definite arc lying in I and having only Z in common with α.

Suppose α_1 is an arc from O to A lying wholly in $O + I$. From here on in this paper the word point will be restricted to mean point of $O + I$ and, accordingly, all the arcs and other point sets under consideration will be understood to be subsets of $O + I$.

Definition: The statement that the arc α_1 takes no step toward spiralling down on A means that it does not cross the straight line interval α. The statement that it takes one step toward spiralling down on A means that it crosses α. The statement that it takes two steps toward spiralling down on A means that there does not exist an arc from O to A which crosses neither α_1 nor α. If n is a positive integer, the statement that α_1 takes $n + 1$ steps toward spiralling down on A means that there do not exist n arcs $\alpha_2, \ldots, \alpha_{n+1}$ from O to A such that if $k < n + 1$, α_{k+1} does not cross α_k, and α_{n+1} does not cross α. The statement that α_1 spirals down on A means that if n is a positive integer, α_1 takes n steps toward spiralling down on A.[1]

It is clear that the following theorem holds true.

THEOREM 1. *If two arcs with O and A as end-points do not cross each other and one of them takes n steps toward spiralling down on A, then, if n is greater than 1, the other one takes $n - 1$ steps toward spiralling down on A and therefore if one of them spirals down on A so does the other one.*

It is to be observed that if an arc takes $n + 1$ steps toward spiralling down on a point then it takes n steps toward spiralling down on that point.

Suppose BXC is an interval of the arc α_1 such that (1) C follows B in the order from O to A on α_1, and (2) there exist a point B' preceding B and a point C' following C, in that order, such that the intervals $B'BX$ and $C'CX$ of α_1 both cross α, and BC is the only subinterval x of the interval $B'C'$ of α_1 which has only its end-points in common with α and contains two arcs, one abutting on α from one side at one end-point of x and the other abutting on α from the other side at the other end-point of x. Then the segment $BXC - (B + C)$ is called a *winding* of α_1 about A and this winding is called positive or negative according as it is or is not true that the interval BX of BXC and the arc LZ abut on α from the same side. No two windings of α_1 about A have a point in common and if two of them have

a common end-point they are either both positive or both negative. It may be shown that if all the windings of α_1 about A are positive or all of them are negative and the total number of them is the positive number n then α_1 takes $n + 1$ steps, and no more, toward spiralling down on A. Let β denote a finite or infinite simple sequence such that (1) each term of β is a collection such that either every element of it is a positive winding of α_1 about A or every element of it is a negative winding of α_1 about A; (2) each winding of α_1 about A is an element of one and only one term of β; (3) if x and y are consecutive terms of β and each element of x is positive then each element of y is negative; and (4) if one collection follows another one in the sequence β then every element of it follows every element of the other one in the order from O to A on the arc α_1.

Suppose no collection of the sequence β has infinitely many elements. Let γ_1 denote a sequence $b_{11}, \ldots$ whose nth term b_{1n} is the total number of windings in the nth term of β or minus that number according as every such winding is positive or every such winding is negative. Let γ_2' denote a sequence whose nth term is $b_{1n} - 1$ or $-(|b_{1n}| - 1)$ according as b_{1n} is positive or negative. If every term of γ_2' is 0 then α_1 takes two steps, and no more, toward spiralling down on A. If there are infinitely many terms of γ_2', all of the same sign, such that no one of them is followed by any term of γ_2' with the opposite sign then α_1 spirals down on A. Suppose that there do not exist infinitely many such terms of γ_2' and that some term of γ_2' is different from 0. Then, and only then, there exists a subsequence γ_2'' of γ_2' such that (1) every two consecutive terms of γ_2'' have different signs; (2) if y is a term of γ_2' lying, in γ_2', between two consecutive terms x and z of γ_2'' and having the same sign as x then there is no term of γ_2' lying, in γ_2', between x and y and having a sign opposite to that of x; and (3) every term of γ_2' different from 0 either is a term of γ_2'' or lies, in γ_2', between some two consecutive terms of γ_2''.

Let γ_2 denote a sequence such that (1) its nth term exists if and only if the nth term of γ_2'' exists, and (2) if x is the nth term of γ_2'' the nth term of γ_2 is x, plus the sum of all other terms y of γ_2' with the same sign as x, if there are any, such that y is not separated from x in γ_2' by any term of γ_2' with a sign opposite to that of x.

Let γ_3' denote a sequence obtained from γ_2 as γ_2' was obtained from γ_1. If every term of γ_3' is 0 then α_1 takes 3 steps, and no more, toward spiralling down on A. If there are infinitely many terms of γ_3', all of the same sign, such that no one of them is followed by any one of the opposite sign then α_1 spirals down on A. If there are not infinitely many such terms of γ_3' and some term of γ_3' is different from 0 then, and only then, there exists a subsequence γ_3'' of γ_3' satisfying with reference to γ_3' the conditions described above as being satisfied by γ_2'' with reference to γ_2' and there exists a sequence γ_3 satisfying with reference to γ_3' and γ_3'' the conditions described above as being satisfied by γ_2 with reference to γ_2' and γ_2''.

There exists a finite or infinite sequence $\gamma_1, \gamma_2', \gamma_2'', \gamma_2, \gamma_3', \gamma_3'', \gamma_3, \ldots$ such that if γ_n' exists then (1) if every one of its terms is 0, α_1 takes n steps, and no more toward spiralling down on A; (2) if there is an infinite subsequence of γ_n', with terms all of the same sign, such that no one of them is followed by any term of γ_n' of the opposite sign then α_1 spirals down on A; and (3) if no such infinite subsequence of γ_n' exists and some term of γ_n' is not 0 then γ_n'', γ_n and γ_{n+1}' exist and relations described above as holding between γ_1', γ_1'', γ_1 and γ_2' continue to hold if they are replaced by γ_n', γ_n'', γ_n and γ_{n+1}', respectively.

The following theorems hold true.

THEOREM 2. *If n is an integer greater than 1 the arc α_1 takes n, and no more, steps toward spiralling down on A if, and only if, γ_n' exists and each of its terms is 0.*

THEOREM 3. *The arc α_1 spirals down on A if, and only if, either (1) some collection of the sequence β has infinitely many elements or (2) for some n, there is an infinite subsequence δ of γ_n' such that all the terms of δ have the same sign and no term of γ_n' of the opposite sign follows any term of δ, or (3) the sequence $\gamma_1, \gamma_2, \gamma_3, \ldots$ is infinite.*

Definition: If the arc OB is a proper subset of the arc OC, the arc OC is said to take n steps toward spiralling down on B if and only if OB does so and OC is said to spiral down on B if and only if OB does so. If the arc XY does not have O as an end-point the statement that it spirals down on the point Z means that it contains Z and is a subset of some arc which has O as an end-point and spirals down on Z. An arc is said to be a *spiral* if it spirals down on some point and a pointset that contains no spiral is said to be *spiral free*.

Definition: If the point C does not belong to the arc OB the arc OB is said to take n steps toward spiralling down on C if and only if every arc from O to C that contains OB takes n steps toward spiralling down on C.

The following theorem holds true.

THEOREM 4. *If the point P is a limit point of the subset M of the arc α_1 and the arc α_1 does not take n steps toward spiralling down on any point of M then it does not take n steps toward spiralling down on P.*

THEOREM 5. *If there are points on which the arc α_1 spirals down the set of all points on which it spirals down is an inner limiting set and every interval of α_1 contains uncountably many points on which α_1 does not spiral down.*

Proof: Let K denote the set of all points on which α_1 does not spiral down. There exists a number m such that there is some point on which α_1 does not take m steps toward spiralling down. For each positive integer n let K_n denote the set of all points P of K such that α_1 does not take $m + n$ steps towards spiralling down on P. The point set K is the sum of the point sets $K_1, K_2, K_3, \ldots$ and, by Theorem 3, each point set of this sequence is closed. It follows that $\alpha_1 - K$ is an inner limiting set.

Suppose XYZ is an interval of α_1. If there is a straight interval containing both O and XYZ then α_1 does not spiral down on any point of XYZ. Suppose no straight interval contains both of them. Then there are two points E and F of XYZ such that O, E and F are the vertices of a triangle OEF. For every point C on the side EF of this triangle let $T(C)$ denote the first point of XYZ, in the order from O, on the straight ray OC. Unless $T(C)$ is X or Z, $T(C)$ is not a limit point of $\alpha_1 - XYZ$ and therefore there is, on the straight interval from O to $T(C)$, a point F such that the straight interval from $T(C)$ to F contains no point of $\alpha_1 - XYZ$ and thus no point of α_1 except $T(C)$ and therefore α_1 does not spiral down on $T(C)$.

Hence, in any case, there are uncountably many points of the interval XYZ on which α_1 does not spiral down.

THEOREM 6. *If M is a compact totally disconnected closed point set not containing O there exists an arc from O which spirals down on each point of M but on no point not belonging to M.*

Proof for the Case Where M is Perfect: There exists a sequence G_1, G_2, G_3, . . . such that (1) for each n, G_n properly covers M and is a collection of 2^n simple domains whose closures are mutually exclusive; (2) each domain of G_n contains the closures of two domains of G_{n+1}; and (3) M is the common part of the point sets G_1^*, G_2^*, G_3^*, ...[2] Let D_1 and D_2 denote the domains of G_1 and let D_{i1} and D_{i2} $(i = 1, 2)$ denote the domains of G_2 that lie in D_i. For each sequence $j_1, j_2, \ldots j_n$ of numbers equal to 1 or 2, $D_{j_1 j_2 \ldots j_n 1}$ and $D_{j_1 j_2 \ldots j_n 2}$ denote the domains of G_{n+1} that lie in the domain $D_{j_1 j_2 \ldots j_n}$ of G_n.

Let J_o denote a triangle OB_1B_2 having O as one of its vertices and neither containing nor enclosing any point of $\bar{G}_1^*$. Let P_1, P_2, P_3, . . . denote a sequence of points between B_1 and B_2, on the side B_1B_2 of this triangle, such that every point of B_1B_2 is a limit point of the set of all points of this sequence. Let O_o denote the point P_1. Let b_1 and b_2 denote the straight intervals B_1O_o and B_2O_o, respectively. Let J_1 and J_2 denote two spiral free simple closed curves containing b_1 and b_2, respectively, having only the point O_o in common, lying, except for b_1 and b_2, wholly in the exterior of J_0, containing no point of $\bar{G}_1^*$, but containing two arcs b_1' and b_2', respectively, such that (1) neither of them intersects b_1 or b_2, and (2) if P is a point of $\bar{D}_1$ $(i = 1, 2)$ and t is an arc lying wholly in J_i plus its interior except that its end-points belong to b_i and b_i', respectively, then t takes one step toward spiralling down on P. There exist two spiral-free simple closed curves J_1' and J_2' containing b_1' and b_2', respectively, lying, except for these arcs, wholly without J_1, J_2 and each other and containing no point of $\bar{G}_2^*$ but such that J_i' $(i = 1, 2)$ contains an arc $B_{i1}B_{i2}$ lying wholly in D_i such that if B_i, B_{i1}, B_{i2} and O_o lie in that order on the boundary of the complementary domain of $J_i + J_i'$ that contains O, and I_i' denotes the interior of J_i', then $\bar{I}_i' \cdot D_i$ and $D_i - D_i \cdot \bar{I}_i'$ are both connected. Let Q_o denote the

collection of all straight intervals from O to points of the interval B_1B_2. Let Q_i $(i = 1, 2)$ denote a collection of mutually exclusive spiral free arcs, filling up J_i plus J_i' plus their interiors I_i and I_i', such that if q is an arc of this collection its end-points are $q \cdot b_i$ and $q \cdot (B_{i1}B_{i2})$. Let O_i denote the point of $B_{i1}B_{i2}$ which is an end-point of the arc of the collection Q_i whose other end-point is the first point of the sequence $P_1, P_2, \ldots$ which belongs to b_i, let b_{i1} $(i = 1, 2)$ denote the interval $B_{i1}O_i$ of $B_{i1}B_{i2}$ and let b_{i2} denote the interval $B_{i2}O_i$ of $B_{i1}B_{i2}$. There exist four spiral-free simple closed curves J_{11}, J_{12}, J_{21} and J_{22} such that (1) no one of them intersects G_2^* or contains a point on or within any other one of them or of J_0, J_1, J_2, J_1' or J_2' except that J_{ij} contains b_{ij}, (2) J_{ij} $(i, j = 1, 2)$ lies wholly in D_i and contains an arc b_{ij}', not intersecting b_{ij}, such that if P is a point of $\bar{D}_{ij}$ and t is an arc lying wholly in J_{ij} plus its interior except that its end-points belong to b_{ij} and b_{ij}', respectively, then t takes 2 steps toward spiralling down on P. There exist four spiral-free simple closed curves J_{11}', J_{12}', J_{21}' and J_{22}' containing b_{11}', b_{12}', b_{21}' and b_{22}', respectively, lying except for these arcs wholly without J_{11}, J_{12}, J_{21}, J_{22} and each other and containing no point of G^* but such that J_{ij}' lies in D_i and contains an arc $B_{ij1}B_{ij2}$, lying wholly in D_{ij}, such that B_{ij}, B_{ij1}, B_{ij2} and O_i lie in that order on the boundary of the complementary domain of $J_{ij} + J_{ij}'$ that contains O and if I_{ij}' denotes the interior of J_{ij}' then $\bar{I}_{ij}' \cdot D_{ij}$ and $D_{ij} - D_{ij} \cdot \bar{I}_{ij}'$ are both connected. Let Q_{ij} $(i, j = 1, 2)$ denote a collection of mutually exclusive spiral-free arcs, filling up $J_{ij} + I_{ij} + J_{ij}' + I_{ij}'$, such that if q is an arc of this collection its end-points are $q \cdot b_{ij}$ and $q \cdot (B_{ij1}B_{ij2})$.

For each point X of $B_{ij1}B_{ij2}$ let $T(X)$ denote the point of B_1B_2 which is an end-point of the arc of the collection Q_i whose other end-point is an end-point of an arc of the collection Q_{ij} whose other end-point is X. Let O_{ij} denote the point of $B_{ij1}B_{ij2}$ which is an end-point of the arc of the collection Q_{ij} whose other end-point is an end-point of an arc of the collection Q_i which has as *its* other end-point the point of lowest subscript in the sequence $P_1, P_2, \ldots$, which lies on b_i between $T(B_{ij1})$ and $T(B_{ij2})$. Continue this process indefinitely. Suppose $k_1, k_2, \ldots k_n, k_{n+1}$ is a sequence of numbers each equal to 1 or 2. The arc $b'_{k_1k_2 \cdots k_n}$ is the common part of the simple closed curves $J_{k_1k_2 \cdots k_n}$ and $J'_{k_1k_2 \cdots k_n}$. If $I_{k_1k_2 \cdots k_n}$ and $I'_{k_1k_2 \cdots k_n}$ denote the interiors of $J_{k_1k_2 \cdots k_n}$ and $J'_{k_1k_2 \cdots k_n}$, respectively, the points $B_{k_1k_2 \cdots k_n}$, $B_{k_1k_2 \cdots k_n1}$, $B_{k_1k_2 \cdots k_n2}$ and $O_{k_1k_2 \cdots k_{n-1}}$ lie in that order on the boundary of $J_{k_1k_2 \cdots k_n} + J'_{k_1k_2 \cdots k_n} + b'_{k_1k_2 \cdots k_n}$, $Q_{k_1k_2 \cdots k_n}$ is a collection of mutually exclusive spiral free arcs filling up $J_{k_1k_2 \cdots k_n} + I_{k_1k_2 \cdots k_n} + J'_{k_1k_2 \cdots k_n} + I'_{k_1k_2 \cdots k_n}$ such that if q is an arc of this collection its end-points are $q \cdot b_{k_1k_2 \cdots k_n}$ and $q \cdot (B_{k_1k_2 \cdots k_n1}B_{k_1k_2 \cdots k_n2})$. For each point X of the arc $B_{k_1k_2 \cdots k_n1}B_{k_1k_2 \cdots k_n2}$, $T(X)$ denotes a point of b_{k_i} for which there exists a sequence of arcs $t_1, t_2, \ldots t_n$ such that for each i $(1 \leqq i \leqq n)$ t_i is an arc of the collection $Q_{k_1 \cdots k_i}$ and every two consecutive arcs of this

sequence have a common end-point and $T(X)$ is an end-point of t_1 and X is an end-point of t_n. The point $O_{k_1k_2\ldots k_n}$ is a point of $B_{k_1k_2\ldots k_n1}B_{k_1k_2\ldots k_n2}$ such that $T(O_{k_1k_2\ldots k_n})$ is the point of lowest subscript in the sequence P_1, P_2, . . . which lies on b_{k_1} between $T(B_{k_1k_2\ldots k_n1})$ and $T(B_{k_1k_2\ldots k_n2})$. The intervals $B_{k_1k_2\ldots k_n1}O_{k_1k_2\ldots k_n}$ and $B_{k_1k_2\ldots k_n2}O_{k_1k_2\ldots k_n}$ are denoted by $b_{k_1k_2\ldots k_n1}$ and $b_{k_1k_2\ldots k_n2}$, respectively. The simple closed curves $J_{k_1k_2\ldots k_n1}$ and $J_{k_1k_2\ldots k_n2}$ have only $O_{k_1k_2\ldots k_n}$ in common and they intersect $J_{k_1k_2\ldots k_n}$ in $b_{k_1k_2\ldots k_n1}$ and $b_{k_1k_2\ldots k_n2}$, respectively. The simple closed curve $J_{k_1k_2\ldots k_nk_{n+1}}$ lies wholly in the domain $D_{k_1k_2\ldots k_n}$ and so does $J'_{k_1k_2\ldots k_nk_{n+1}}$, and if P is a point of $\bar{D}_{k_1k_2\ldots k_nk_{n+1}}$ every arc of the collection $Q_{k_1k_2\ldots k_nk_{n+1}}$ takes n steps toward spiralling down on P.

Let Z denote a collection of simple disks such that z belongs to Z if, and only if, either z is the triangle OB_1B_2 plus its interior or, for some finite sequence $k_1, \ldots k_n$ each term of which is 1 or 2, Z is $J_{k_1\ldots k_n} + I_{k_1\ldots k_n}$ or $J'_{k_1\ldots k_n} + I'_{k_1\ldots k_n}$. The sum of all the disks of the collection Z is a disk whose boundary J is a simple closed curve containing M. The arc OB_1B_2 on J spirals down on every point of M and on no point that does not belong to M.

THEOREM 7. *If OB_1B_2 is a triangle there exists a collection G of arcs such that (1) each of them has O as one of its end-points and no two of them have any point in common except O, (2) every straight interval with one end-point at O and the other one on the side B_1B_2 of this triangle is a subset of some arc of the collection G and each arc of the collection G contains one such interval, (3) each arc of the collection G spirals down on some point, and (4) the set of all points P such that some arc of G spirals down on P is a perfect point set.*

Proof: Using the notation of the above proof of a special case of Theorem 5, let Q denote a collection such that q is an element of it if and only if there exists a finite sequence $k_1, \ldots, k_n$, of which each element is 1 or 2, such that q is an arc of the collection $Q_{k_1k_2\ldots k_n}$. Suppose X is a point of the straight interval B_1B_2 not belonging to the sequence $P_1, P_2, \ldots$. Then there is only one subcollection $Q(X)$ of Q such that X plus the sum of the arcs of $Q(X)$ plus some point of K is an arc from X to that point of K. Let $T(X)$ denote that point of K. If X is a point of the sequence $P_1, P_2, \ldots$ there are only two such collections, $Q_1(X)$ and $Q_2(X)$. In this case, for each number i which is equal either to 1 or to 2, let $T_i(X)$ denote a point of K such that $X + T_i(X) + Q_i^*(X)$ is an arc with X and $T_i(X)$ as its end-points.

There exists a sequence of mutually exclusive spiral-free arcs $\alpha_1, \alpha_2, \ldots$ such that, for each n, α_n has $T_1(P_n)$ and $T_2(P_n)$ as its end-points and lies, except for those points, wholly in the exterior of J. Let G' denote a collection such that g belongs to it if, and only if, either (1) g is $X + Q^*(X) + T(X)$ for some point X of B_1B_2 not belonging to $P_1, P_2, \ldots$, or (2) for some n, g is $P_n + Q_1^*(P_n) + \alpha_n$. Let G denote a collection of arcs such that the arc g belongs to it if, and only if, it is an arc of G' plus some straight interval

with one end-point at O. The collection G fulfills all the requirements of the conclusion of Theorem 6.

[1] Professor Eilenberg has indicated to me that Waraszkiewicz has considered certain "spirals" on a cone which bear to one another a relationship similar to the relationship that α_1 bears to α according to this definition. Cf. Waraszkiewicz, S., *Fundamenta Mathematicae*, **22**, 180–205 (1934) and *Ibid.*, **23**, 172–189 (1934).

[2] The notation G^* is used to denote the sum of all the point sets of the collection G.

HOMOLOGY GROUPS OF SYMMETRIC GROUPS AND REDUCED POWER OPERATIONS

BY N. E. STEENROD

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated by S. Lefschetz, January 13, 1953

1. *Introduction.*—We shall give a generalization of the reduced power operations[1] which connects them with the homology groups $H_i(S_n)$ of the symmetric group S_n of degree n. Briefly if u is a q-dimensional cohomology class of a complex X, and $c \in H_i(S_n)$, then we define a cohomology class $u^n/c \in H^{nq-i}(X)$ called the nth power reduced by c. These operations include those defined previously. The broader viewpoint gives natural explanations of relations among cyclic reduced powers found by Adem[2] and Thom.[3] Their relations are homology relations in S_n. The new definition was inspired by the work of Adem.

2. *Definition of Reduced Powers.*—Let X^n denote the product complex of n copies of the complex X. An element u of the group $C^q(X; B)$, of q-cochains of X with coefficient group B, has an nth power defined by

$$u^n \cdot (\sigma_1 \times \ldots \times \sigma_n) = (u \cdot \sigma_1) \otimes \ldots \otimes (u \cdot \sigma_n) \qquad (2.1)$$

where σ_i is a cell of X and $u \cdot \sigma_i \in B$ is the value of u on σ_i. Thus $u^n \in C^{nq}(X^n; B^n)$ where B^n is the tensor product of n copies of B.

Let π be a permutation group of the factors of X^n. Each $\alpha \in \pi$ gives an automorphism of the integral chain groups $C_r(X^n)$. For example, if α interchanges the first two factors, then

$$\alpha(\sigma_1 \times \sigma_2 \times \ldots \times \sigma_n) = (-1)^{rs}\sigma_2 \times \sigma_1 \times \sigma_3 \times \ldots \times \sigma_n \qquad (2.2)$$

where r, s are the dimensions of σ_1, σ_2. We require u^n to be an *equivariant* cochain, i.e.,

$$u^n \cdot \alpha(\sigma_1 \times \ldots \times \sigma_n) = \alpha(u^n \cdot \sigma_1 \times \ldots \times \sigma_n). \qquad (2.3)$$

This and 2.1, 2.2 determine the operation of α in B^n, namely, if q is even, α permutes the factors of B^n in the same manner as the factors of X^n; and,

if q is odd, it effects the same permutation but multiplied by the sign of the permutation. The symbol $B^{n,q}$ denotes the π-group B^n with the operators chosen as above depending on the parity of q. The equivariant cochain groups are denoted by ${}^\pi C^r(X^n)$. Thus $u^n \in {}^\pi C^{nq}(X^n; B^{n,q})$.

Let W be an acyclic complex on which π operates freely, i.e., for each cell e of W the cells αe, as α ranges over π, are pairwise distinct. The construction of such a W for any group π is a basic step[4] in defining the homology groups of π. Let A be a coefficient group with π as left operator group. Let $\alpha \in \pi$ operate in $A \otimes W$ by $\alpha(a \otimes e) = (\alpha a) \otimes (\alpha e)$. A chain of $A \otimes W$ which is a sum of chains of the form $\alpha c - c$ is called *residual*. These chains form a subcomplex; the corresponding factor complex is denoted by $A \otimes_\pi W$. Its homology groups are independent of the choice of W and are written $H_i(\pi; A)$.

In $W \times X$ let $\alpha \in \pi$ operate by

$$\alpha(e \times \sigma) = (\alpha e) \times \sigma. \tag{2.4}$$

Define the diagonal carrier ϕ: $W \times X \to X^n$ by $\phi(e \times \sigma) = \sigma^n$. Then ϕ is equivariant ($\phi\alpha = \alpha\phi$) and acyclic. Since W is π-free, so also is $W \times X$. Therefore it is possible to choose an equivariant chain mapping

$$\phi_\#: W \times X \to X^n, \qquad \phi_\#\alpha = \alpha\phi_\# \tag{2.5}$$

carried by ϕ. This is accomplished by the usual construction by induction on the dimension with the modification that, once we have defined $\phi_\#(e \times \sigma)$ for all σ's and one e of an equivalence class $\{\alpha e\}$, we then set $\phi_\#(\alpha e \times \sigma) = \alpha\phi_\#(e \times \sigma)$. Since $\alpha \neq \beta$ implies $\alpha e \neq \beta e$, the procedure is valid.

Since $\phi_\#$ is equivariant, its induced cochain mapping $\phi^\#$ preserves equivariance. Thus $\phi^\# u^n$ is an equivariant cochain of $W \times X$.

We define now a reduction process passing from cochains of $W \times X$ to cochains of X. This is done without restriction on W, X. Let $v \in C^r(W \times X; G)$ and let $c = \sum a_j e_j$ be an i-chain of W with coefficients in A. If σ is an $(r - i)$-cell of X, then $c \times \sigma = \sum a_j(e_j \times \sigma)$ is an r-chain of $W \times X$. Define $v \cdot (c \times \sigma) \in A \otimes G$ by

$$v \cdot (c \times \sigma) = \sum_j a_j \otimes (v \cdot e_j \times \sigma). \tag{2.6}$$

This is a mild extension of the Kronecker index, and continues to satisfy

$$\delta v \cdot (c \times \sigma) = v \cdot \partial(c \times \sigma). \tag{2.7}$$

Now define the $(r - i)$-cochain v/c of X by

$$(v/c) \cdot \sigma = v \cdot (c \times \sigma). \tag{2.8}$$

The symbol is read "v slant c" and is called the *reduction* of v by c. It has coefficients in $A \otimes G$. Using the standard formula for $\partial(c \times \sigma)$ and 2.7, we obtain

$$\delta v/c = v/\partial c + (-1)^i \delta(v/c). \tag{2.9}$$

We are thus led to a bilinear operation v/c defined for a cohomology class v and a homology class c. It is the reduction process associated with the $\times$-product of homology classes.[5]

With the earlier hypotheses on W, suppose v is an equivariant r-cochain of $W \times X$ with coefficients in the π-group G. For any $\alpha \in \pi$, and chain c of W, we have

$$[v/(\alpha c - c)] \cdot \sigma = v \cdot [(\alpha c - c) \times \sigma] = (\alpha - 1)(v \cdot c \times \sigma).$$

Thus $v/(\alpha c - c)$ has residual values in $A \otimes G$. Factoring by residual elements gives $A \otimes_\pi G$, and let η be the cochain map induced by the natural map $A \otimes G \to A \otimes_\pi G$. Then

$$\eta(v/(\alpha c - c)) = 0. \tag{2.10}$$

Extend the definition of v/c to the case of an equivariant v and a $c \in C_i(A \otimes_\pi W)$ by choosing a representative c' of c in $A \otimes W$ and setting

$$v/c = \eta(v/c'). \tag{2.11}$$

Then $v/c \in C^{r-i}(X; A \otimes_\pi G)$. The formula 2.9 continues to hold for the extended slant operation.

Now set $v = \phi^\# u^n$ so that $G = B^{n,q}$. Then

$$\phi^\# u^n/c \in C^{nq-i}(X; A \otimes_\pi B^{n,q}) \tag{2.12}$$

is called the reduction by c of the nth power of u (abbreviated: a reduced nth power of u). Let u be a cocycle, then so also are u^n and $\phi^\# u^n$. Taking c to be a cycle, it follows from 2.9 that $\phi^\# u^n/c$ is a cocycle. If the cycle c is varied by a boundary then 2.9 shows that $\phi^\# u^n/c$ varies by a coboundary. If u is varied by a coboundary, then u^n also varies by a coboundary. However the variation is not generally the coboundary of an *equivariant* cochain. Because of this a more complicated argument is required to show that $\phi^\# u^n/c$ varies by a coboundary. This argument is a moderate generalization of the one already given for the earlier definition.[1] Thus 2.12 defines an operation for a cohomology class $u \in H^q(X; B)$, and a homology class $c \in H_i(\pi; A)$, and the value is denoted by

$$u^n/c \in H^{nq-i}(X; A \otimes_\pi B^{n,q}). \tag{2.13}$$

We drop the ϕ since it is easily seen that any two choices of ϕ_π are equivariantly homotopic.

Whenever $i > nq - q$ and σ is an $(nq - i)$-cell, then the chain $\phi_\pi(c \times \sigma)$ has the dimension nq which exceeds the dimension of its carrier σ^n. Therefore $\phi_\pi(c \times \sigma) = 0$, and we have

$$u^n/c = 0 \quad \text{if} \quad i > nq - q. \tag{2.14}$$

If a cocycle u of X is zero on a subcomplex X_0', it is easily seen that $\phi^i u^n/c$ is also zero on X_0. It follows that u^n/c is defined for elements u in the *relative* cohomology groups of (X, X_0), and u^n/c is likewise relative. Everything below extends in a natural way to the relative case.

The connection between the present and former definitions of reduced powers is the following. Let α_i $(i = 1, 2, \ldots)$ be a 0-sequence[1] in the group ring $Z(\pi)$. Let c_0 be a vertex of W. Then $In(\alpha_1 c_0) = 0$ since α_1 has 0 as its coefficient sum. As W is acyclic, it contains a 1-chain c_1 with $\partial c_1 = \alpha_1 c_0$. Then

$$\partial(\alpha_2 c_1) = \alpha_2 \partial c_1 = \alpha_2 \alpha_1 c_0 = 0, \qquad \text{since} \quad \alpha_2 \alpha_1 = 0.$$

As W is acyclic, it contains a c_2 with $\partial c_2 = \alpha_2 c_1$. Continue inductively and construct a sequence c_i such that $\partial c_i = \alpha_i c_{i-1}$. If W is reduced by residual chains, c_i becomes a cycle modulo the coefficient sum of α_i. It is easy now to identify $\phi^i u^n/c_i$ with the $D_i u^n$ in the earlier definition.

3. *Invariance Properties.*—If f is a map $X \to Y$, and f^* is the induced homomorphism of cohomology groups of Y into those of X, and $u \in H^q(Y; B)$, then

$$f^*(u^n/c) = (f^* u)^n/c. \tag{3.1}$$

This embodies topological invariance: take f to be a homeomorphism.

Suppose $\pi \subset \pi'$ are two subgroups of the symmetric group S_n. Then the inclusion map $\lambda: \pi \to \pi'$ induces two maps

$$\lambda_*: H_i(\pi; A) \to H_i(\pi'; A), \quad \lambda': A \otimes_\pi B^{n,q} \to A \otimes_{\pi'} B^{n,q},$$

where A is a π'-module. Clearly λ' induces homomorphisms λ' of the cohomology groups of X with coefficients in $A \otimes_\pi B^{n,q}$. Then we have, for $u \in H^q(X; B)$ and $c \in H_i(\pi; A)$,

$$u^n/\lambda_* c = \lambda'(u^n/c). \tag{3.2}$$

In those cases where λ' is the identity, this becomes

$$u^n/c = u^n/\lambda_* c. \tag{3.3}$$

This is true in the simplest cases. For example, take $A = B$ to be a cyclic group of finite or infinite order. Then $B^n \approx B$. If q is even, then S_n operates as the identity in A; and if q is odd, it operates by multiplying by the sign of the permutation. It follows that $A \otimes_S B^{n,q} \approx A$. Thus the same cyclic coefficient group occurs throughout. In these cases we have the following principle: *For ease of calculation, choose π to be the smallest subgroup of S_n "containing" the cycle c; and to obtain relations on u^n/c, choose $\pi = S_n$.*

Let Z, Z_m denote the integers and integers mod m. Let

$$\partial_*: H_i(\pi; Z_m) \to H_{i-1}(\pi; Z); \quad \delta^*: H^s(X; Z_m) \to H^{s+1}(X; Z)$$

be the boundary, coboundary operators associated with the exact coefficient sequence

$$0 \to Z \xrightarrow{\xi} Z \xrightarrow{\eta} Z_m \to 0$$

where $\xi(k) = mk$ and η is the natural factorization $Z/\xi Z$. Using 2.9, we obtain immediately

$$u^n/\partial_* c = (-1)^{i+1}\delta^*(u^n/c), \qquad c \in H_i(\pi;\, Z_m), \qquad u \in H^q(X;\, Z). \quad (3.4)$$

There are two additional invariance properties corresponding to homomorphisms of the coefficients A or B. Their statements are obvious.

If $\pi = 1$, then $H_i(\pi;\, Z) = 0$ for $i > 0$, and $H_0(\pi;\, Z) \approx Z$ has a generator represented by a vertex c_0 of W. Clearly $u^n/c_0 \in H^{qn}(X;\, B^n)$. Now $\phi_\# c_0 \times X$ defines a chain map approximating the diagonal map $X \to X^n$. Therefore

$$u^n/c_0 \text{ is the } n\text{th power of } u \text{ in the cup-product sense.} \quad (3.5)$$

[1] Steenrod, N. E., *Ann. Math.*, **56**, 47–67 (1952).

[2] Adem, J., these PROCEEDINGS, **38**, 720–726 (1952).

[3] Thom, R., Strasbourg Colloq., 1951 (mimeographed).

[4] Eilenberg, S., and MacLane, S., *Ann. Math.*, **48**, 51–78 (1947).

[5] There is another reduction which corresponds to the cohomology cross-product. It yields the cap-product under a diagonal approximation.

CYCLIC REDUCED POWERS OF COHOMOLOGY CLASSES

BY N. E. STEENROD

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY

Communicated by S. Lefschetz, January 13, 1953

In a previous paper[1] we have shown that each homology class c of the symmetric group S_n provides a topologically invariant operation defined for each cohomology class u of a complex, and it gives a cohomology class u^n/c called the nth power reduced by c. In this paper we state properties of the simplest of these, namely, those provided by cycles c lying on the cyclic subgroup of S_n of order n. We shall continue with the notation and numbering of the previous paper.

4. Homology Groups of the Cyclic Group.—Let π be the cyclic group of permutations of the factors of X^n, and T its generator which increases the index of the factor by 1 mod n. In the group ring $Z(\pi)$, set

$$\Delta = T - 1, \quad \Sigma = \sum_{k=0}^{n-1} T^k. \quad (4.1)$$

Construct a π-complex W having one cell e_i and its transforms $T^k e_i$ in each dimension $i \geqq 0$. Define ∂ by

$$\partial e_{2i+1} = \Delta e_{2i}, \quad \partial e_{2i+2} = \Sigma e_{2i+1}, \quad i \geqq 0. \tag{4.2}$$

Clearly W is π-free. Also W is acyclic; for each of Δ, Σ generates the ideal of annihilators of the other.

We shall use for coefficients the integers Z and the integers mod n, Z_n; and we suppose π operates trivially. Now $Z \otimes_\pi W = W/\pi$ is W factored by residual chains. Hence $C_j(W/\pi)$ is an infinite cyclic group generated by e_j; and the boundary relations obtained from 4.2 are

$$\partial e_{2i+1} = 0, \quad \partial e_{2i+2} = n e_{2i+1}, \qquad i \geqq 0. \tag{4.3}$$

Therefore

$$H_{2i+1}(\pi; Z) \approx Z_n; \quad H_{2i+2}(\pi; Z) = 0, \qquad i \geqq 0, \tag{4.4}$$

$$H_j(\pi, Z_n) \approx Z_n, \qquad j \geqq 0. \tag{4.5}$$

In each case that $H_j \neq 0$, it is a cyclic group generated by the cycle e_j. We let e_j represent also its homology class.

If ∂_* is defined as in §3 with $m = n$, and e_{2i} generates $H_{2i}(\pi; Z_n)$, then 4.3 shows that $\partial_* e_{2i} = e_{2i-1}$ in $H_{2i-1}(\pi; Z)$, and $\partial_* e_{2i}$ mod n is e_{2i-1} in $H_{2i-1}(\pi; Z_n)$. In view of 3.4, it follows that the reduced powers u^n/e_{2i} determine u^n/c for c in $H_{2i-1}(\pi)$ with coefficients Z or Z_n. We may therefore restrict future considerations to $e_{2i} \in H_{2i}(\pi; Z_n)$.

Thom[2] has proved the following result:

4.6. THEOREM. *If n is an odd prime, and $q = \dim u$, then $u^n/e_{2i} = 0$ whenever the change in dimension, namely $(nq - 2i) - q$, is not an even multiple of $n - 1$.*

A new proof of this is obtained by considering the homomorphisms

$$h_{2i}: H_{2i}(\pi; Z_n) \to H_{2i}(S_n; Z_n) \tag{4.7}$$

induced by the inclusion map of π in the symmetric group S_n. We must distinguish two cases of 4.7:

(1°) each $\alpha \in S_n$ operates as the identity in Z_n,

(2°) each $\alpha \in S_n$ operates in Z_n by the sign of α.

Since n is odd, π contains only even permutations; hence, in both cases π operates trivially in Z_n. The results to be proved are:

4.8. THEOREM. *In case 1°, $h_{2i} = 0$, if $2i$ is not an even multiple of $n - 1$. In case 2°, $h_{2i} = 0$ if $2i$ is not an odd multiple of $n - 1$.*

We obtain Thom's result by combining 4.8 with 3.3 using case 1° or 2° according as q is even or odd.

To prove 4.8, let k be prime to n. Define $\gamma \in S_n$ by $\gamma(s) = ks \bmod n$ $(s = 0, 1, \ldots, n - 1)$. Let ψ be the inner automorphism of S_n provided

by γ. It is readily verified that $\psi T = T^k$; hence $\lambda = \psi | \pi$ is an automorphism of π. Now λ, ψ induce automorphisms λ_*, ψ_* of $H(\pi)$ and $H(S_n)$. Clearly $h_{2i}\lambda_* = \psi\, h_{2i}$. Since ψ is an *inner* automorphism, ψ_* is the identity. Therefore $\lambda_* e_{2i} - e_{2i}$ lies in the kernel of h_{2i}.

To compute λ_*, we define a chain map $\lambda_\sharp$: $W \to W$ such that $\lambda_\sharp(\alpha c) = (\lambda\alpha)(\lambda_\sharp c)$ by

$$\lambda_\sharp e_{2i} = k^i e_{2i}, \quad \lambda_\sharp e_{2i+1} = k^i \sum_{m=0}^{k-1} T^m e_{2i+1}. \tag{4.9}$$

One verifies readily that $\partial\lambda_\sharp = \lambda_\sharp\partial$. Therefore $\lambda_\sharp$ induces the required λ_*. Now let k be a primitive root of the prime n. Then γ is an odd permutation; and so, in case 2°, ψ and λ reverse signs in Z_n. Therefore

$$\lambda_* e_{2i} - e_{2i} = \begin{cases} (k^i - 1)e_{2i}, & \text{case 1°}, \\ -(k^i + 1)e_{2i}, & \text{case 2°}. \end{cases} \tag{4.10}$$

Since $ne_{2i} = 0$, we seek the G. C. D. of n and $k^i - 1$ in case 1°, and of n and $k^i + 1$ in case 2°. As k is a primitive root, n divides $k^i - 1$ only when i is a multiple of $n - 1$; and n divides $k^i + 1$ only when i is an odd multiple of $(n-1)/2$. In all other cases the G. C. D. is 1 so that e_{2i} belongs to the kernel of h_{2i}.

5. Product Formulas for Cyclic Reduced Powers.—Cartan[3] has proved a formula for reduced squares which may be written in the notation of §4 as

$$(u \times v)^2/e_i = \sum_{j=0}^{i} (u^2/e_j) \times (v^2/e_{i-j}) \bmod 2 \tag{5.1}$$

where

$$u \in H^q(X),\ v \in H^r(Y), \text{ and } u \times v \in H^{q+r}(X \times Y).$$

Taking $X = Y$ and using the diagonal map $X \to X \times X$, one obtains the same formula with cup-products in place of cross-products. We shall give a generalization of this formula to cyclic reduced nth powers.

Let π be the cyclic group of order n, and let d: $\pi \to \pi \times \pi$ be the diagonal map. Using the complex W for π given in §4, we define a chain map $d_\sharp$: $W \to W \times W$ by

$$d_\sharp e_{2i} = \sum_{j=0}^{i} e_{2j} \times e_{2i-2j} + \sum_{j=0}^{i-1} \Omega e_{2j+1} \times e_{2i-2j-1}$$

$$d_\sharp e_{2i+1} = \sum_{j=0}^{i} \{e_{2j} \times e_{2i-2j+1} + e_{2j+1} \times Te_{2i-2j}\}. \tag{5.2}$$

Here $\Omega = \sum T^k \times T^l$ summed over the range $0 \leqq k < l \leqq n-1$. As usual we set

$$d_\sharp T^m e_j = (dT^m) d_\sharp e_j = T^m \times T^m d_\sharp e_j.$$

It can be verified that $\partial d_\sharp = d_\sharp \partial$ and $d_\sharp(Tc) = (dT)d_\sharp c$. Therefore the homomorphism $H_j(\pi) \to H_j(\pi \times \pi)$ induced by $d_\sharp$ is the d_* induced by d. Reducing modulo residual chains, the map $d_\sharp$ of W/π into $(W/\pi) \times (W/\pi)$ becomes

$$d_\sharp e_{2i} = \sum_{j=0}^{i} e_{2j} \times e_{2i-2j} + \frac{1}{2} n(n-1) \sum_{j=0}^{i-1} e_{2j+1} \times e_{2i-2j-1}$$
$$d_\sharp e_{2i+1} = \sum_{j=0}^{i} \{e_{2j} \times e_{2i-2j+1} + e_{2j+1} \times e_{2i-2j}\}. \tag{5.3}$$

From these equations, d_* may be determined in the various cases. For example, if $n = 2$ with coefficients Z_2 we have

$$d_* e_i = \sum_{j=0}^{i} e_j \times e_{i-j}. \tag{5.4}$$

Again, if n is odd with coefficients Z_n, we have

$$d_* e_{2i} = \sum_{j=0}^{i} e_{2j} \times e_{2i-2j}, \qquad \text{mod } n,$$
$$d_* e_{2i+1} = \sum_{s=0}^{2i+1} e_s \times e_{2i+1-s}, \qquad \text{mod } n. \tag{5.5}$$

To apply these formulas to reduced powers, we suppose that π operates as the cyclic permutation on the factors of both X^n and Y^n. Then $\pi \times \pi$ operates in $X^n \times Y^n$. Let λ, μ be the natural chain maps

$$\lambda: X^n \times Y^n \to (X \times Y)^n, \quad \mu: W \times W \times X \times Y \to W \times X \times W \times Y.$$

Let $\phi_\sharp$: $W \times X \to X^n$, and $\psi_\sharp$: $W \times Y \to Y^n$ be equivariant maps belonging to the diagonal carrier (see §2). Set

$$\theta_\sharp = \lambda(\phi_\sharp \times \psi_\sharp)\mu: W \times W \times X \times Y \to (X \times Y)^n.$$

Suppose $u \in C^q(X)$, $v \in C^r(Y)$, then a short calculation gives

$$\theta^\sharp(u \times v)^n / e_s \times e_t = \epsilon\{\phi^\sharp u^n / e_s\} \times \{\psi^\sharp v^n / e_t\} \tag{5.6}$$

where ϵ is -1 to the exponent $t(nq - s) + qrn(n-1)/2$. Let d' : $W \times X \times Y \to W \times W \times X \times Y$ be the cross-product map of d with the identity of $X \times Y$. Then

$$\theta_\sharp d'_\sharp: W \times X \times Y \to (X \times Y)^n$$

is equivariant and belongs to the diagonal carrier; hence it may be used to compute reduced nth powers. Obviously

$$d'^\sharp \theta^\sharp (u \times v)^n / e = \theta^\sharp (u \times v)^n / d_\sharp e. \tag{5.7}$$

Substitute here the formula 5.3 for $d_\sharp e$, use the linearity of the slant opera-

tion, and then expand each term according to 5.6. The result is a formula expressing a reduced power of $u \times v$ as a sum of products of reduced powers of u and v. Passing to cohomology classes, we obtain 5.1 when $n = 2$. If n is odd with coefficients Z_n, we obtain as in 5.5

$$(u \times v)^n/e_{2i} = \epsilon \sum_{j=0}^{i} (u^n/e_{2j}) \times (v^n/e_{2i-2j}) \tag{5.8}$$

where ϵ is -1 to the exponent $qr(n-1)/2$. If n is prime, then 4.8 shows that we need consider only those indices $2i$, $2j$ in 5.8 which are even or odd multiples of $n-1$ depending on the dimensions of u, v.

6. Computations of Special Cases.—We give first two formulas for cyclic reduced nth powers where n is odd. Setting $m = (n-1)/2$, these are

$$u^n/e_{q(n-1)} = (-1)^{mq(q-1)/2}(m!)^q u, \tag{6.1}$$

$$(\delta u)^n/e_{i+n-1} = (-1)^{mq} m!\, \delta(u^n/e_{2i}). \tag{6.2}$$

Both relations hold mod n, and u is a q-cohomology class with coefficients Z or Z_n. In the second, u is a cohomology class of a subcomplex A of X, δ is the natural homomorphism $H^r(A) \to H^{r+1}(X, A)$, and i is unrestricted.

In passing from a space X to its suspension X' there is a suspension isomorphism s: $H^r(X) \approx H^{r+1}(X')$. Now s is a composition of a single δ and several isomorphisms induced by maps.[4] Since induced homomorphisms commute with reduced powers, it follows that 6.2 holds with s in place of δ.

When $q = 1$, 6.1 becomes

$$u^n/e_{n-1} = m!u. \tag{6.3}$$

First, we will show how 6.1, 2 are proved assuming that 6.3 holds in the special case that u is the generator of $H^1(I, S)$ where $I = [0, 1]$ is a 1-cell and S is its boundary consisting of the vertices 0 and 1. Now

$$(I, S) \times (I, S) = (I \times I, I \times S \cup S \times I)$$

is a 2-cell and its boundary and $u \times u$ generates its H^2. Apply 5.8 with $u = v$ and $i = n-1$; then the sum has one non-zero term which is evaluated by 6.3. Thus 6.1 holds for the 2-class $u \times u$. Continue inductively with $(I, S)^q = (I, S)^{q-1} \times (I, S)$ and $u^{(q)} = u^{(q-1)} \times u$. It follows that 6.1 holds for the generating class of a q-cell modulo its boundary. Shrinking the boundary to a point, we obtain the same for a q-sphere modulo a point. Now any q-class on a q-dimensional complex is representable as the image of the generating class of the q-sphere under a continuous map. By invariance under maps, 6.1 follows for any q-class on a q-complex. For any complex X the inclusion map of the q-skeleton $X^q \subset X$ induces an isomorphic imbedding $H^q(X) \to H^q(X^q)$. This gives the general case of 6.1.

To obtain 6.2 form the product $(I, S) \times A$. Since $S \times A$ consists of two disjoint copies of A, identifying A with $1 \times A$ maps $H^q(A)$ isomorphically into a direct summand of $H^q(S \times A)$. Let $u' \in H^q(S \times A)$ correspond to u. A generator v of $H^1(I, S)$ can be selected so that $\delta u' = v \times u \in H^{q+1}(I \times A, S \times A)$. Applying 5.8 and 6.3 we obtain

$$(\delta u')^n/e_{2i+n-1} = (v \times u)^n/e_{2i+n-1} = (-1)^{mq}(v^n/e_{n-1}) \times (u^n/\epsilon_{2i})$$
$$= (-1)^{mq} m!\, v \times (u^n/e_{2i}) = (-1)^{mq} m!\, \delta(u'^n/e_{2i}).$$

Thus 6.2 holds for u'. Let X' be the space obtained by pinching to a point the subset $0 \times A$ of $I \times A$. Clearly $H^r(X', 1 \times A) \to H^r(I \times A, S \times A)$ has kernel zero for all r. Thus 6.1 holds for the δ of $(X', 1 \times A)$. Since X' is contractible to a point, the map $A \to 1 \times A$ extends to a map f: $(X, A) \to (X', 1 \times A)$. Since f^* commutes with δ and reduced powers, 6.1 follows for the δ of (X, A).

The validity of 6.3 for the generator of $H^1(I, S)$ is established by a direct calculation of a chain map $\phi_\sharp$: $W \times I \to I^n$. Letting U, V represent the vertices 0, 1, we set

$$\begin{aligned}
\phi_\sharp(e_0 \times U) &= U^n, \quad \phi_\sharp(e_j \times U) = 0 \text{ for } j > 0,\\
\phi_\sharp(e_0 \times V) &= V^n, \quad \phi_\sharp(e_j \times V) = 0 \text{ for } j > 0,\\
\phi_\sharp(e_{2i} \times I) &= i! \sum_{a,b} U^{a_0} I V^{b_0} I U^{a_1} I V^{b_1} \ldots I U^{a_i} I V^{b_i}, \qquad (6.4)\\
\phi_\sharp(e_{2i+1} \times I) &= i! \sum_{a,b} I U^{a_0} I V^{b_0} I U^{a_1} I V^{b_1} \ldots I U^{a_i} I V^{b_i}.
\end{aligned}$$

The exponents a, b are $\geqq 0$, and the sum of the a's and b's is n-$2i$-1 in the first case, and n-$2i$-2 in the second. In both cases the sum is taken over all possible distinct choices of the a's and b's. The verification of $\partial\phi_\sharp = \phi_\sharp\partial$ is a direct counting process. Then $\phi_\sharp(e_{n-1} \times I) = m!\, I^n$ and 6.3 follows.

In applying 6.1, 2, it is useful to keep in mind the following facts about $m!$ from elementary number theory:

$$\text{if } n \text{ is composite and } n \neq 9, \text{ then } m! \equiv 0 \bmod n, \qquad (6.5)$$

$$\text{if } n \text{ is prime and } m \text{ is even, then } (m!)^2 \equiv -1 \bmod n, \qquad (6.6)$$

$$\text{if } n \text{ is prime and } m \text{ is odd, then } m! \equiv \pm 1 \bmod n. \qquad (6.7)$$

Thus when n is an odd prime, the coefficients in 6.1, 2 are non-zero mod n. These awkward coefficients can be eliminated in the following manner, suggested to me by Serre. When n is an odd prime, define

$$\mathcal{P}^s u = (-1)^{ms+mq(q-1)/2}(m!)^{2s-q} u^n / e_{(q-2s)(n-1)}, \qquad (6.8)$$

where $q = \dim u$ and u is arbitrary. Then, for any complex, $\mathcal{P}^s$ is a homomorphism

$$\mathcal{P}^s\colon H^q(X; Z_n) \to H^{q+2s(n-1)}(X; Z_n)$$

satisfying the following conditions:

$$\mathcal{P}^0 = \text{identity}, \tag{6.9}$$

$$\text{when } q \text{ is even, } \mathcal{P}^{q/2}u = u^n \text{ (in cup-product sense)}, \tag{6.10}$$

$$\text{when } s > q/2, \ \mathcal{P}^s u = 0, \tag{6.11}$$

$$\mathcal{P}^s(u \times v) = \sum_{t=0}^{s} \mathcal{P}^t u \times \mathcal{P}^{s-t} v, \tag{6.12}$$

$$6.12 \text{ holds with cup-products in place of cross-products,} \tag{6.13}$$

$$\mathcal{P}^s f^* = f^* \mathcal{P}^s \tag{6.14}$$

$$\mathcal{P}^s \delta = \delta \mathcal{P}^s \tag{6.15}$$

$$6.15 \text{ holds with } \delta \text{ replaced by a suspension isomorphism.} \tag{6.16}$$

All relations hold modulo the prime n. 6.9 is a restatement of 6.1, 6.10 of 3.5, 6.12 and 6.13 of 5.8, 6.14 of 3.1, and 6.15 and 6.16 of 6.2; and 6.11 is a convention which is consistent with 6.8 if we agree that $e_j = 0$ for $j < 0$.

Notice that 6.9, 10, and 11 determine $\mathcal{P}^s u$ for all 1 and 2-dimensional u. Combining with 6.13 we can determine $\mathcal{P}^s u$ in the part of the cohomology ring of a complex generated by 1- and 2-dimensional classes. For example, if u is 2-dimensional and u^k is its kth power in the sense of cup-products, then a simple induction based on 6.13 gives

$$\mathcal{P}^s(u^k) = \binom{k}{s} u^{k+s(n-1)} \qquad \text{mod } n, \tag{6.17}$$

where ( ) denotes the binomial coefficient. This formula determines $\mathcal{P}^s$ completely in the complex projective space of arbitrary dimension since its cohomology ring is the polynomial ring in a single 2-dimensional class.

[1] Steenrod, N. E., these PROCEEDINGS, **39**, 213–217 (1953).
[2] Thom, R., Strasbourg Colloq., 1951 (mimeographed).
[3] Cartan, H., *C. R. Acad. Paris*, **230**, 425–427 (1950).
[4] Steenrod, N. E., *Ann. Math.*, **48**, 290–320 (1947), see §11.

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The PROCEEDINGS will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be comprehensive in giving the relation of the paper to previous publications of the author or of others and in exhibiting, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work and long tables of data should be avoided, but authors should be precise in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though more elaborate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

REFERENCES to literature, numbered consecutively, will be placed at the end of the article and short footnotes should be avoided. It is suggested that references to periodicals be furnished in some detail and in general in accordance with the standard adopted for the Subject Catalogue of the International Catalog of Scientific Literature, viz., name of author, with initials following (ordinarily omitting title of paper), abbreviated name of Journal, with place of publication, series (if any), volume, inclusive pages, year. For example: Montgomery, T. H., *J. Morph.*, Boston, 22, 731–815 (1911); or, Wheeler, W. M., *Konigsberg, Schr. physik. Ges.*, 25, 1–142 (1914).

MANUSCRIPTS will be accepted only from members of the Academy or from chairmen of the divisions of the National Research Council who will assume responsibility for the propriety and scientific standards of the paper and should be addressed to *Proceedings of the* NATIONAL ACADEMY OF SCIENCES, *2101 Constitution Ave., Washington, 25, D. C. Articles of which a member of the Academy is author or a co-author may extend to twelve printed pages, those by non-members only to six printed pages; space above these limits will be billed to the author at $17.50 per page.*

PROOF will ordinarily be sent,—with the understanding that charges for an author's correction shall be billed to him. Authors are therefore requested to make final revisions on the typewritten manuscripts.

REPRINTS should be ordered at the time of submission of manuscript. They will be furnished to authors at cost, approximately as follows:

SCHEDULE OF RATES FOR REPRINTS, CARRIAGE CHARGES EXTRA

No. Copies	50	100	150	200	250
4 pp.	4.55	5.20	5.85	6.50	7.15
8 pp.	9.10	10.40	11.70	13.00	14.30
Covers	3.45	4.43	5.41	6.39	7.37

NOTE: These prices will be increased by 8% effective April 1, 1953.

CONTENTS

Page

ASTRONOMY.—NEW WHITE DWARFS IN THE SOUTHERN HEMISPHERE: THIRD LIST . . . . . *By Willem J. Luyten* 135

CHEMISTRY.—A KINETIC TREATMENT OF SATURATION AND CONDENSATION STATISTICS . . . . . *By Henry Eyring and Merrill Wallenstein* 138

CHEMISTRY.—THE HIGH FIELD CONDUCTANCE OF SOME SYMMETRICAL VALENCE-TYPE STRONG ELECTROLYTES . . . . . *By Andrew Patterson, Jr.* 146

GENETICS.—THE BASIS OF THE OXYGEN EFFECT ON X-IRRADIATED DROSOPHILA SPERM . . . . . *By William K. Baker and Elisabeth S. Von Halle* 152

GENETICS.—THE GENETICS OF HOMEOSTASIS IN DROSOPHILA . . . . . *By Th. Dobzhansky and Bruce Wallace* 162

GENETICS.—HETEROGENEITY OF CLONES OF SACCHAROMYCES DERIVED FROM HAPLOID ASCOSPORES . . . . . *By Herschel Roman and Stanley M. Sands* 171

MATHEMATICS.—ASSUMPTION OF BOUNDARY VALUES AND THE GREEN'S FUNCTION IN THE DIRICHLET PROBLEM FOR THE GENERAL LINEAR ELLIPTIC EQUATION . . . . . *By Felix E. Browder* 179

MATHEMATICS.—LINEAR PARABOLIC DIFFERENTIAL EQUATIONS OF ARBITRARY ORDER; GENERAL BOUNDARY-VALUE PROBLEMS FOR ELLIPTIC EQUATIONS. . . . . *By Felix E. Browder* 185

MATHEMATICS.—STOCHASTIC PROCESSES WITH FINITE AND NON-FINITE VARIANCE . . . . . *By S. Bochner* 190

MATHEMATICS.—A PROPERTY OF MINIMAL SURFACES. . . . . *By Robert Finn* 197

MATHEMATICS.—SECOND ORDER ELLIPTIC SYSTEMS OF DIFFERENTIAL EQUATIONS . . . . . *By Charles B. Morrey, Jr.* 201

MATHEMATICS.—SPIRALS IN THE PLANE. . . . . *By R. L. Moore* 207

MATHEMATICS.—HOMOLOGY GROUPS OF SYMMETRIC GROUPS AND REDUCED POWER OPERATIONS . . . . . *By N. E. Steenrod* 213

MATHEMATICS.—CYCLIC REDUCED POWERS OF COHOMOLOGY CLASSES . . . . . *By N. E. Steenrod* 217